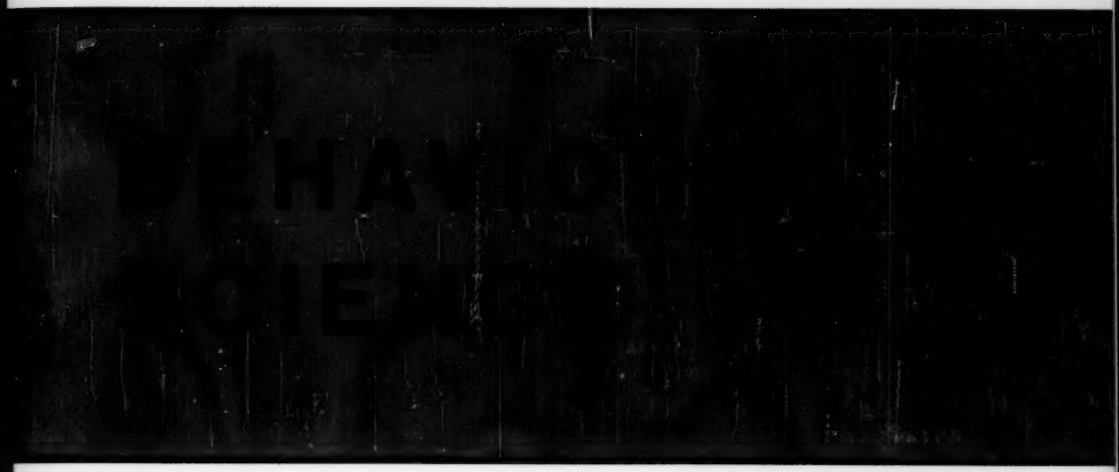
BF

THE UNIVERSITY OF MICHIGAN

VOLUME 6 • NUMBER 1 • JANUARY 1961

1961

PERIODICAL READING ROOM



CONTENTS

1 Clyde Kluckhohn (1905–1960) *by George P. Murdock*

5 Computation, Behavior, and Structure in Fixed and Growing Automata *by Arthur W. Burks*

23 Aspects of Judgment and Decision Making: Interrelationships and Changes with Age *by Michael A. Wallach and Nathan Kogan*

37 The Factor Structure of Variables Concerned with Coronary Artery Disease *by Lee D. Cady, Jr., Menard M. Gertler, Lida G. Gottsch, and Max A. Woodbury*

42 Stochastic Models of the Single Case as the Basis of Clinical Research Design *by Jacob B. Chassan*

51 Dedication of the Mental Health Research Institute Building

CRITIQUE AND COMMENT

72 A Note on the Appearance of Widsom in Large Bureaucratic Organizations *by Karl W. Deutsch and William G. Madow*

79 A Committee Report on Schizophrenic Language *by John R. Parks*

84 Book Reviews

87 Harry H. Goode (1909–1960) *by Merrill M. Flood*

COMPUTERS IN BEHAVIORAL SCIENCE

88 Some Problems and Approaches to Automation of Medical Diagnosis *by Philip A. Smith*

91 Suggestions for Self-Adapting Computer Models of Brain Functions *by Leonard Uhr and Charles Vossler*

97 Computer Program Abstracts

99 Abstracts of Current Literature with Interdisciplinary Implications

102 About the Authors

BEHAVIORAL SCIENCE

A quarterly journal appearing in January, April, July, and October
Official Publication of the Mental Health Research Institute, The University of Michigan. Published at Mt. Royal & Guilford Aves., Baltimore 2, Maryland and the Mental Health Research Institute, The University of Michigan, Ann Arbor, Michigan

Behavioral Science contains articles on general theories of behavior and on empirical research specifically oriented toward such theories. An interdisciplinary approach to problems of behavior is stressed. Although the scope of the journal will include all aspects of behavior which can be subsumed under broadly general interdisciplinary theory, in the field of application special emphasis is placed on contributions relating to research in mental health and disease.

EDITORIAL BOARD: *Franz Alexander*, psychoanalysis—*Alex Bavelas*, social psychology—*Karl W. Deutsch*, political science—*David Easton*, political science—*Ralph W. Gerard*, biology (neurophysiology)—*Marion J. Levy, Jr.*, sociology—*Donald G. Marquis*, psychology—*Jacob Marschak*, economics—*James G. Miller*, psychiatry and psychology—*George P. Murdock*, anthropology—*Theodore M. Newcomb*, social psychology—*Anatol Rapoport*, mathematical biology—*Ralph W. Tyler*, education—*Raymond W. Waggoner*, psychiatry

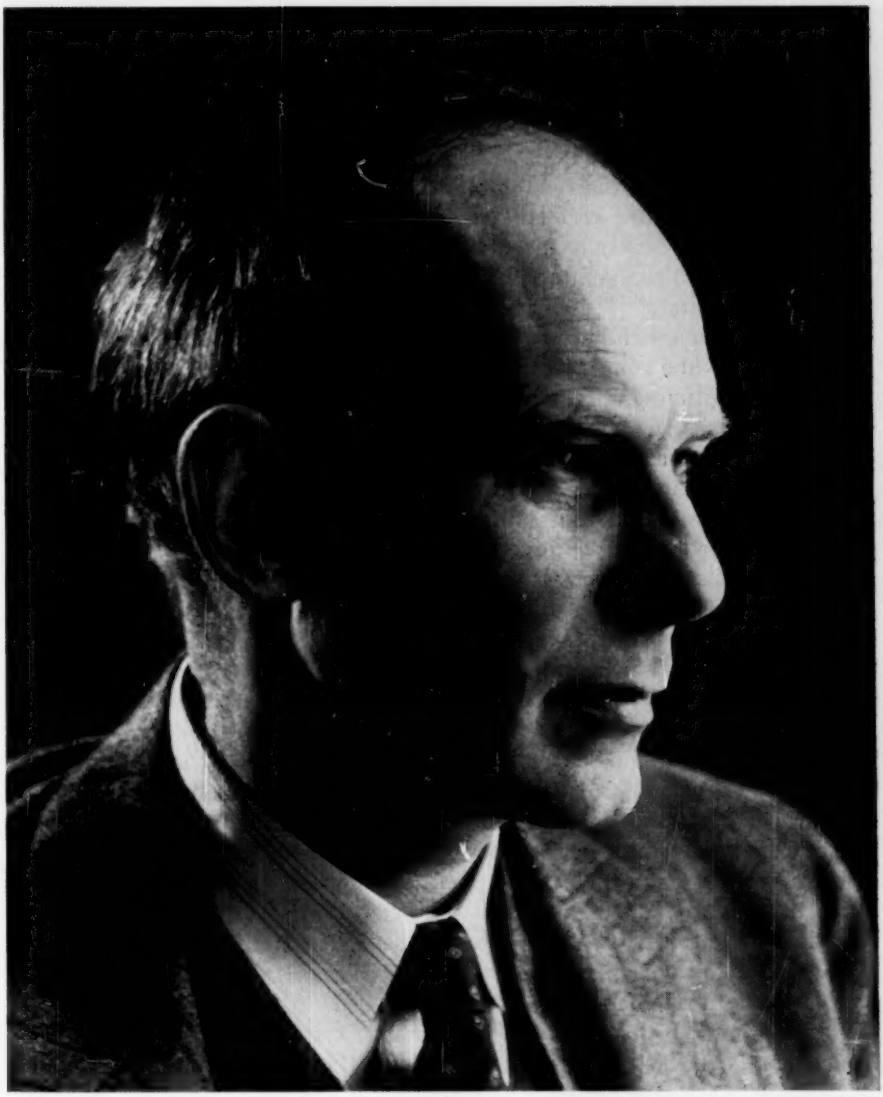
Manuscripts should be sent in care of Dr. James G. Miller, Mental Health Research Institute, The University of Michigan, Ann Arbor, Michigan

COMPUTERS IN BEHAVIORAL SCIENCE EDITORIAL COMMITTEE: *Steven G. Vandenberg*, University of Louisville—*Bert F. Green*, Massachusetts Institute of Technology—*Bruce MacKenzie*, International Business Machines—*Charles Wrigley*, Michigan State University.

Manuscripts for the Computers in Behavioral Science department should be sent in care of Dr. Steven G. Vandenberg, Mental Health Research Institute, The University of Michigan, Ann Arbor, Michigan.

Subscription price: $6.00 a year. Price per issue: $1.75
Send subscriptions, address changes, and all business communications to *Behavioral Science*, Mental Health Research Institute, The University of Michigan, Ann Arbor, Michigan. Make all checks payable to *Behavioral Science.*

Second-class postage paid at Baltimore, Md.
Copyright, 1961, by the Mental Health Research Institute, The University of Michigan.



Clyde Kay Maben Kluckhohn

CLYDE KLUCKHOHN (1905-1960)

Clyde Kay Maben Kluckhohn was born at Le Mars, Iowa, on January 11, 1905, the son of George Wesley and Katherine Swanzey Kluckhohn. Following an elementary education in the public schools, he attended Culver Military Academy (1918–21) and the Lawrenceville School (1921–22). He entered Princeton in the fall of 1922, but the aftermath of an attack of rheumatic fever suffered in childhood forced him to defer his college education. Through a family connection he retired to the Vogt ranch in Ramah, New Mexico, to recuperate his health. This event was largely to shape his subsequent career.

While at the ranch he spent much of his time on pack trips through the adjacent Navaho country. His experiences stimulated him to write *To the Foot of the Rainbow* (1927), followed by *Beyond the Rainbow* (1933). These sensitive books, though addressed to a literary audience, reveal his love of the country and his appreciation of its indigenous inhabitants, and prefigure his choice of a professional career in anthropology. After his recovery, he revisited the Southwest nearly every summer when he was not abroad or engaged in government work in Washington. It was during the last of these trips that he succumbed to a heart attack in Santa Fe on July 29, 1960.

Instead of returning to Princeton, Kluckhohn transferred to the University of Wisconsin in order to be near his adoptive parents (after the death of his mother, he had been adopted by a maternal uncle). At Wisconsin he was elected to Phi Beta Kappa and received the A.B. degree in 1928. A Rhodes Scholarship then sent him to Corpus Christi College at Oxford. He spent the year 1931–32 in Vienna where he began the formal study of anthropology at the University of Vienna and underwent eight months of psychoanalysis. Returning to the United States, he married Florence Rockwood on October 15, 1932, and was appointed Assistant Professor of Anthropology at the University of New Mexico (1932-34). After the birth of his only child, Richard, he spent seven additional months at Oxford and then enrolled as a graduate student at Harvard, where he was to spend the balance of his career.

At Harvard, Kluckhohn received the Ph.D. degree in 1936 and served successively as Instructor in Anthropology (1935-37), Assistant Professor (1937-40), Associate Professor (1940-46), and Professor (1946-60). In 1937 he and Florence Kluckhohn, with Henry A. Murray, Talcott Parsons, and Robert Waelder, formed an interdisciplinary discussion group, called the Levellers, which eventuated in the establishment of the Department of Social Relations at Harvard. Among his other university activities, of comparable importance was his founding of the Russian Research Center there in 1947. This organization, which he served as director

until 1954, has made an imposing contribution to substantive knowledge of the Soviet Union.

During World War II, Kluckhohn moved to Washington where he held civilian positions with the War Department and the Office of War Information. In 1946-47 he acted as an expert consultant to General MacArthur's headquarters, and in 1956 he became a member of an advisory committee to the Foreign Service Institute of the Department of State.

Kluckhohn was a superb teacher. He believed firmly in undergraduate education and in the desirability of exposing future lawyers and business men to the broadening perspectives of anthropology. It was, however, with graduate students and younger colleagues that he excelled. He opened his home to them, showed a warm interest in their work and their ideas, and was generous with sound and realistic advice. Quick to detect quality, he rarely expressed praise directly, but those whom he appreciated or valued soon learned of his good opinion through others to whom he had spoken or through some convincing evidence of material support.

As an ethnographer Kluckhohn devoted most of his life to the Navaho. In addition to two general books of a descriptive and interpretive nature, *The Navaho* (1946) and *Children of the People* (1947), both written in collaboration with Dorothea Leighton, he described particular aspects of their culture in a number of articles and monographs, of which his *Navaho Witchcraft* (1944) is generally regarded by his professional colleagues as his most original and profound contribution. Through a series of major research projects which he initiated at Harvard, he introduced a very large number of students and others to field work among the Navaho and to comparative research among other groups in the Southwest, notably the Zuñi Indians and communities of Mexicans, Mormons, and Texans.

Kluckhohn's theoretical interests converged most strongly on the concept of culture, on the emergent field of culture and personality, and on the problem of values, reflected respectively in *Culture: A Critical Review of Concepts and Definitions* (1952) written in collaboration with A. L. Kroeber, in *Personality in Nature, Society, and Culture* (1948) edited in co-operation with Henry A. Murray, and in a series of more recent papers. His prize-winning *Mirror for Man* (1949) interpreted anthropology in illuminating fashion to the educated layman and to specialists in other fields.

Kluckhohn's wide-ranging interests frequently carried him abroad, e.g., to Mexico in 1946, to Australia in 1953, to Japan in 1954, and to India in 1957. Wherever he went, he made friends—for anthropology and his country as well as for himself. He spoke fluent French and Spanish, as well as German and Navaho. While testifying in a land claims case the year before his death, the prosecution asked him whether he could speak Navaho. When he modestly replied, "A little, but not very well," the Navaho tribal council, who were attending the session, broke courtroom etiquette by shouting in unison, "But he does! He does!"

His scientific contributions and his personal effectiveness brought Kluckhohn a multitude of honors. He was a member of the American Academy of Arts and Sciences and of the National Academy of Sciences, serving on its Committee on Behavioral Sciences. He was Chairman of the Division of Anthropology and Psychology of the National Research Council in 1956-58. The University of New Mexico awarded him an honorary L.H.D. degree in 1949. He was elected President of the American Anthropological Association in 1947, and received the Viking Fund medal and award in general anthropology in 1950.

As suggested by the breadth of his interests and by his competence in such adjacent fields as psychology and linguistics, Kluckhohn was probably the most intellectually sophisticated anthropologist of his generation. Though always loyal to his own discipline, he encouraged exploration and promoted every sound effort toward the integration of the behavioral sciences (including this journal, through active service on its Board of Editors). Though ambitious, he lent his support to students and colleagues whose work he adjudged to be creative with a generosity unparalleled among his peers. If Margaret Mead typifies anthropology to the general public, Clyde Kluckhohn represented the subject to his professional colleagues. He was *par excellence* the behavioral scientists' anthropologist, and the void he has left will not soon be filled.

GEORGE PETER MURDOCK
University of Pittsburgh

The theory of automata is relevant to behavioral science in the sense that it is an absolutely rigorous theory of processes which are themselves absolutely rigorous abstractions of behavior. If we can specify exactly how we wish a system to behave (i.e., what shall happen in response to what), we can in principle construct the system. To what extent, then, can "higher order" systems of this sort be constructed, systems which will construct other systems with prescribed behavior characteristics? In particular, a question which has intrigued automata theoreticians is whether an automaton can be constructed which will reproduce itself and, if so, how complex does it need to be? These questions and related matters are discussed in this article.

COMPUTATION, BEHAVIOR, AND STRUCTURE IN FIXED AND GROWING AUTOMATA

by Arthur W. Burks

The University of Michigan

1. INTRODUCTION[1]

THIS paper is concerned mainly with deterministic rather than probabilistic automata. It is perhaps true that deterministic theories are inadequate to account fully for growth, learning, mutation, evolution, and the actual operation of digital computers, and hence that a probabilistic theory is required for a full explanation of all these phenomena. Nevertheless, we cannot fully understand probabilistic automata until we know the limits of deterministic ones, for only then will we know what can be done by probabilistic automata that cannot be done by deterministic machines. Moreover, a probabilistic automaton may be regarded as a deterministic automaton to which has been added a probability measure governing the transitions between states. Just as both deterministic and probabilistic theories have been important in physical science, we may expect both kinds to be fruitful in the study of automata.

A complete theory of self-organizing systems must show exactly what role determinism plays in these systems. Von Neumann (1956) suggested that automata theory must include probabilistic as well as deterministic logics, and that deterministic analyses of phenomena such as self-reproduction are of interest (see Section 4 below).

Various species of deterministic automata have been studied. Turing machines have been intensively investigated, both directly and through the medium of recursive functions (Kleene, 1952; Davis, 1958). Fixed (finite) automata have received much attention of late, and von Neumann worked some with self-reproducing automata. Little has been done, however, to relate these separate inquiries. The present paper is directed toward a unified theory of deterministic automata.

2. STRUCTURE, BEHAVIOR, AND COMPUTATION IN FIXED AUTOMATA AND GENERALIZED TURING MACHINES

Let us begin with fixed automata and Turing machines and turn later to self-reproducing machines and automata generally. A fixed automaton[2] (Burks & Wright,

[1] This is a revised version of a paper of the same title which appeared in *Self-organizing systems* (edited by Marshall Yovits and Scott Cameron), New York, Pergamon Press, 1960, pp. 282–311. The writing of this paper was supported by the Office of Naval Research, and by the U. S. Army Signal Corps through Project MICHIGAN. John Holland, Richard Laing, and Jesse Wright have made many helpful suggestions.

[2] Strictly speaking, these automata should be called *finite* fixed automata, since there are also

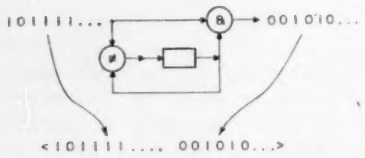


FIG. 1. Behavior of fixed automaton.

1953; Burks & Wang, 1957) consists of a finite number of switch and delay elements interconnected so as to make a pattern or structure which is the same at every moment of time. Figure 1 represents a very simple fixed automaton which deletes every other pulse (1) that is fed into the input. A digital computer with a finite tape is, of course, a fixed automaton. (A digital computer with an indefinitely expandable tape is a generalized Turing machine in the sense defined below.)

The structure of a fixed automaton is to be distinguished from its behavior. The *structure* consists in the fixed pattern of interconnections of the elements. The structures of two automata are the same if the nodes (i.e., the junctions of the wires of the elements) of these automata can be placed in one-one correspondence so that corresponding nodes are nodes of identical elements.[3] Thus if we change the conjunction of Figure 1 into a disjunction we obtain a different structure.

The behavior of an automaton consists, roughly speaking, of the relation of input stimuli to output responses. A more precise definition of behavior may be given in terms of temporal sequences of states, where the relevant times are the nonnegative integers 0, 1, 2, 3. . . . Let us call an infinite temporal sequence of input states to an automaton an *input history* and an infinite temporal sequence of output states an *output history*. A *behavior pair* of an automaton consists of an input history paired with the output history which is produced when that input history is impressed on the inputs of the automaton. A behavior pair for the automaton of Figure 1 is shown in the lower part of the figure. The *behavior* of a fixed automaton is the set of all behavior pairs of that automaton. It should be kept in mind that what we call "histories" are *infinite* sequences of states.

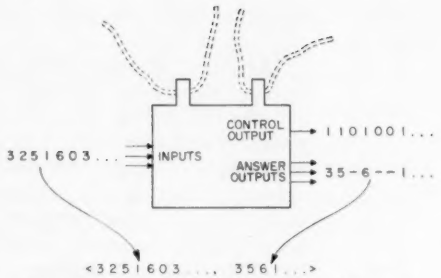


FIG. 2. Computation of generalized Turing machine.

We turn next to Turing machines. As Turing originally conceived them (Turing, 1936), each is essentially a fixed automaton connected to a single infinite tape. The fixed automaton scans one square of the tape at any one time, and can move the tape (or move relatively to the tape) one square at a time. We will generalize Turing's concept in a number of ways to obtain what we call a *generalized Turing machine*. To relate Turing machines to the fixed automata previously discussed, we will allow the fixed part of the machine to have inputs and outputs, as in Figure 2; as a consequence, the earlier concept of *behavior* applies directly to Turing machines. Next we will allow any finite number of tapes. Finally, we will regard a tape at any given moment of time as a finite automaton and construe a change in tape length as a change in this automaton; thus for us the tape is finite at each moment of time.[4] A tape "square" can be synthesized from switch

infinite fixed structures composed of infinitely many switch and delay elements. The latter, however, will be discussed under the heading of "growing automaton," and for convenience we will use "fixed automaton" to mean *finite* fixed automaton.

[3] Elements are taken to include their input and output wires.

[4] There are other ways of treating Turing machines. For example, one can consider an infinitely long tape with a fixed automaton moving along it. Note, however, that though there are an infinite number of squares on this tape, at each moment of time only a finite number of them have ever changed state; it is for this reason that we can treat the tape as finite at each moment of time.

and delay elements in a number of ways, and squares can be added or subtracted to represent the growing or shrinking of the tape.[5] Hence the definition of structural equivalence given above can be extended to take account of tapes. While a generalized Turing machine has a given structure at any moment of time, its structure changes through time in a very simple fashion and under the control of the computation going on in the machine. A digital computer together with an operator who can indefinitely expand the magnetic tapes is a generalized Turing machine.

Computation requires a problem and an answer, and the answer must be presented somewhere. Turing used alternate squares of the tape for this purpose. We will stipulate that the answer of a generalized Turing machine is to appear on the outputs. We will not, however, take the output history as the answer, for a reason to be explained below. Instead, we will distinguish a *control output wire* from the remainder of the output wires, which will be called *answer output wires*. A "1" on the control output wire signifies that the state of the answer output wires at that time is part of the answer, while a "0" on the control output wire means that the state of the answer output wires at that time is to be ignored. The subsequence of the states of the answer output wires so selected is the *computed output sequence*. The sequence 3561 is the first part of the computed output sequence of Figure 2. A pair consisting of an input history and the resultant computed output sequence is a *computation pair*. The *computation* of an automaton is the set of all computation pairs of that automaton.

There is a relational analogy between behavior and computation on the one hand and "real time" simulation and "nonreal time" simulation on the other hand. In real time simulation the answer is produced at a rate determined by the actual time taken by the process which the computer is simulating, while in a "nonreal time" simulation the answer is produced at a rate independent of actual time. Similarly, the behavior of an automaton includes the outputs at every moment of time (including the control output), while the computation includes only outputs selected by the computer itself, and in general these answer states do not appear at a uniform rate.

A computed output sequence may be null, finite, or infinite. A computation is said to be *infinite* if every computed output sequence (i.e., the second element of each computation pair) is infinite, *finite* if every computed output sequence is finite or null, and *mixed* otherwise.

We can apply the concept of computation (and the related concepts just defined) to fixed automata by distinguishing a control output of a fixed automaton from the other (or answer) outputs. Hence all three concepts—structure, behavior, and computation—apply to both fixed automata and generalized Turing machines. These three concepts will also apply to the growing automata defined later (Section 4).

It is easy to see that the class of infinite computations of generalized Turing machines includes the class of behaviors of such machines as a proper part; the same holds true for fixed automata. Consider the behavior of an automaton A; this is the computation of an automaton A' made by taking the outputs of A as the answer outputs of A' and making the control output of A' always active (always in state 1). But there are infinite computations which are not deterministic in character and hence which are not the behavior of any deterministic automaton. An example is the computation $I(0)$, $I(1)$, $I(2), \ldots$; $I(1)$, $I(2)$, $I(3), \ldots$ [where $I(t)$ is the input state at t], which is realized by a very simple fixed automaton: the output state at t is the input state at t, and the control output node C has the behavior $C(0) = 0$, $C(t + 1) = 1$ for all t. Hence (for either fixed automata or generalized Turing machines) the class of infinite computations is broader than but inclusive of the class of behaviors. For this reason we did not take the whole output history of a machine as the answer computed by this machine—to do this would be to iden-

[5] The synthesis of a tape "square" storing one bit (a "0" or a "1") is described in Section 4 below and also in Section 6 of Burks (1959). One can construct tape squares capable of storing any of n characters in a similar way.

tify behavior and computation and hence to restrict unduly the domain of the computable.

3. ANALYSIS AND SYNTHESIS OF AUTOMATA

The discussion of automata in Sections 4 and 5 will presuppose some knowledge of fixed automata and generalized Turing machines, and so we will mention briefly some results about these special cases of automata and will note some of the unsolved problems. Since the concepts of behavior, structure, and computation apply to all automata (see Sections 4 and 5), many of the results and problems concerning fixed automata and generalized Turing machines which we will discuss in the present section may be extended to cover automata generally.

Much is known about the behavior of fixed automata. For example, each fixed automaton has a finite number of delay output states, which fact has important consequences for the behavior of these devices. In the case of a fixed automaton without inputs, it means that the behavior of the automaton is periodic.[6] In the case of an automaton with inputs, the finitude of states means that the device can detect only "regular events" (Copi, Elgot, & Wright, 1958), that if the number of delay states is known, then its behavior can be computed from the results of a finitely long behavioral test, and that there is a decision procedure for behavioral equivalence (Moore, 1956; Burks & Wang, 1957). Since the notion of computation is a new one for fixed automata, little is known about this. The same argument that shows that the behavior of a fixed automaton without inputs is periodic shows that the computation of such an automaton is also periodic. There is an algorithm for deciding whether the computation of a fixed automaton (with inputs) is infinite, finite, or mixed (Burks & Wright, unpub.), but the problem of the existence of an algorithm for deciding whether two fixed automata have the same computation is unsolved.

[6] Burks and Wright (1953), Theorem I. Note that in this case the behavior is a single output history.

A great deal is known about Turing computability and the closely related notions of algorithm, general recursiveness, and partial recursiveness. We do not have space to explain these concepts here (see Turing, 1936; Kleene, 1952; Davis, 1958), but it is worth noting that these results can be expressed in terms of our notion of a generalized Turing machine, and they contain answers to most questions about the computation of such machines. As applied to generalized Turing machines, these results on recursiveness, algorithms, and computability need interpretation, however, because there are many ways of relating a generalized Turing machine to these notions. Thus there are alternative methods of using a generalized Turing machine to compute a recursive function. For example, for a given nonrelative recursive function, one can design a machine without inputs whose computation will be an enumeration of the values of that function. Alternatively, one can design a generalized Turing machine with inputs so that whenever arguments (problems) are presented on the input, the computed functional values (answers) will appear as part of the computation. The distinction between finite or mixed computation on the one hand and infinite computation on the other (see Section 2) can be made to correspond to the distinction between partial recursiveness and general recursiveness.

We will relate a few of the things known about computability and recursiveness to our concept of a generalized Turing machine. Every computable number (in Turing's sense) can be produced as the computed output sequence of a generalized Turing machine with no inputs and one tape. The values of a function which is general recursive relative to a given function can be produced as the computation of a generalized Turing machine which receives the values of the given function as inputs. There is no algorithm for deciding whether the computation of an arbitrary generalized Turing machine is finite, mixed, or infinite, but there is an algorithm for deciding whether the structure of a machine can ever change. There is no algorithm for deciding whether

two generalized Turing machines produce the same computation (Wright, unpub.).

Since the notion of behavior as applied to a Turing machine is new, there are many open questions in this area. For example: is the class of behaviors of generalized Turing machines with $n + 1$ tapes larger than the class of behaviors for machines with n tapes? Is there an algorithm for deciding whether two generalized Turing machines have the same behavior?

Later we will discuss von Neumann's universal constructing machine, so Turing's universal computing machine is of particular interest here. Let us call a generalized Turing machine with no inputs, one tape, one control output node, and one answer output node a special purpose machine. Translated into our terms, Turing's result is that there is a one tape generalized Turing machine with inputs, called a universal Turing machine, with this property: for each special purpose machine M there is a finitely long program or input sequence S such that when S is supplied to the inputs of the universal Turing machine the computed output sequence of M and the universal Turing machine are identical; this is illustrated in Figure 3. It should be noted that in general the universal Turing machine and the special purpose machine do not produce the same output histories, so that the equivalence between them has to do with computation and not behavior. It should be noted also that the finitely long sequence S can be supplied to the universal Turing machine by connecting a cycle free sequence of delays without inputs to each of the inputs of the universal machine; the input to these delays is fed by a contradiction (an output that is always false), and the program is stored in the initial states of the delays. After a period of time equal to the length of this sequence of delays, the input to the universal machine is always zero.

It is very likely that a similar result holds when inputs are added to the special purpose machine and more than one tape is allowed; that is, for each set of generalized Turing machines with n inputs and m outputs, there is a universal Turing machine. Does

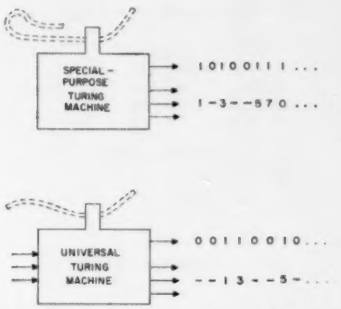


FIG. 3. Universal Turing machine.

an analogous result hold for fixed automata? Computer engineers often speak of trading time for hardware and of the choice between constructing ("wiring in") an operation and instructing (programming) that operation, and this might seem to imply the existence of a universal fixed automaton. But in fact there are no universal fixed automata. We will prove that there is no universal fixed automaton for the class of input-free fixed automata with m binary outputs. To prove this we first show that any fixed automaton A (with inputs) when supplied by a finite input sequence S produces a periodic output of period no greater than n, where n is the number of delay states of A. To say that S is finite is to say that there will come a time when the input states of A are always the same thereafter. The automaton A will then behave as a noninput device with n delay states, and so the computed output sequence of A is periodic with period n or less. Hence A cannot produce a computed output sequence of, for example, period $2n$, and so A is not a universal machine. Thus there is no fixed automaton which is universal for the class of noninput fixed automata with m outputs, and consequently there is no fixed automaton for the class of fixed automata with n inputs and m outputs.

We will next mention some reduction problems. Does adding more tapes increase the class of computations produced by generalized Turing machines? A generalized Turing machine can have only one multichannel reading and writing head per tape. In contrast, the growing automata intro-

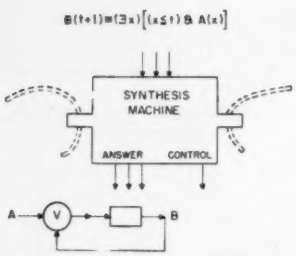


FIG. 4. Synthesis machine.

duced in the next section can have tapes with more than one multi-channel read-write head attached to them.[7] What effect, if any, does this have on the class of behaviors or the class of computations? In some cases one can use several single-headed tapes alternately to do the work of one many-headed tape. This can be done in the special case of an automaton which records information on two tapes and compares these tapes, using one head for recording and another head on the same tape for reading and comparing. To accomplish this result with one-head tapes, we use two such tapes for each of the two-head tapes on the original machine. At each moment one of these tapes is used to record the input information while the other is used for the comparison; when the latter tape is emptied, it is used for recording and the former tape is used for comparing. It is worth noting that we have here an abstract buffering and scheduling problem.

A universal Turing machine can produce the computation of any special purpose Turing machine. From this idea of a machine mimicking another machine, the concept of a machine *constructing* another machine arises. This synthesis process can be conceived in many ways. One method which has been studied some of late (Büchi, Elgot, & Wright, 1958; Church, 1958; Büchi & Elgot, 1958; Elgot, 1959) is illustrated in Figure 4. Suppose we have a formal language L in which a human can conveniently and precisely state some behavioral condition C which he would like an automaton to satisfy. We would like to know if there is a synthesis machine S of which we can prove mathematically the following result: when S is given any behavioral condition C, it will always produce either the design of some automaton A which satisfies this condition, or it will tell us that there is no such automaton.[8] The condition given at the top of Figure 4 is that for all times other than zero the output of the desired automaton is to be active (in state 1) if the input has been active at some previous time. (That is, B is active at time $t + 1$ if and only if there is a time x no larger than t at which A is active.) There are fixed automata satisfying this condition, and so if a synthesis machine for a language containing this condition exists, this machine would produce as output a representation of one of these automata.

There is a large area of research here because there are many problems of the kind just described. The synthesis process represented in Figure 4 starts with a particular language L and results in the synthesis of a fixed automaton satisfying a condition on its behavior. There are many different languages L which would be investigated and it is of interest to synthesize generalized Turing machines as well as fixed automata. In both cases, we can allow the condition to be a condition on the computation of the desired automaton as well as a condition on its behavior. For each specification of the language L, the nature of the condition C, and the nature of the desired automaton A, the question arises: is there a synthesis machine and if so, what is its structure? In some cases there is a machine and in some cases there is not. It is important to realize that, if there is no synthesis machine of this kind, there is no method of using computers to design computers which is guaranteed to produce an answer in all cases. Hence the broad problem is this: to determine the theoretical limits of mechaniz-

[7] This can be done by constructing an indefinite series of storage squares, and moving the reading heads along them. Motion of a reading "head" is accomplished by "destroying" the head and constructing another head in the place where it is wanted.

[8] Other problems in this area are formulated by Büchi, Elgot, and Wright (1958). What we have called a synthesis problem they call a combined solvability and synthesis problem.

ability of the synthesis process by deterministic machines of this kind.

4. GROWING AUTOMATA

John von Neumann (1951) has shown how to design universal constructing machines and self-reproducing automata within the framework of a general definition of automata. (See also Kemeny, 1955; Shannon, 1958.) The details of von Neumann's definition are contained in an unpublished manuscript, and so an alternative definition will be offered here. Von Neumann's model is much closer to biological phenomena than the definition presented here; for example, each of von Neumann's cells has a unit delay (Shannon, 1958, p. 127) while many of our cells have no delays. Von Neumann's constructions are also based on very weak primitives (his cells are capable of only 29 different states [Shannon, 1958]) while we have assumed very strong primitives (each cell may have any of about 28,000 structures and each structure is capable of several states). While the definition presented here is much less realistic and economical than von Neumann's, it does make it easy to construct and present various kinds of growing automata.

A comment about our use of the word "automaton" is needed here. The concept we will define is that of *automaton*, without qualification. The *fixed automata* of Section 2 will be special cases of automata, namely, those automata consisting of fixed structures of computing elements. Automata whose structures change in time are called *growing automata*. But although our definition and some of our subsequent remarks apply to automata generally, we are particularly interested in growing automata.

The basis of each automaton is an infinite two-dimensional array of discrete cells.[9] A particular automaton is defined by specifying the structure of a finite number of cells (i.e., stating what elements are in them) at time zero; the rest of the cells are empty (structureless). The structure at time $t + 1$ is determined by the structure at t and the complete state of the device at t, according to rules to be explained. These rules do not allow more than a finite number of cells to be structured at any particular time.

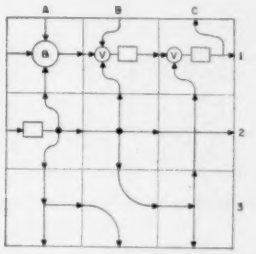


FIG. 5. Computing primitives.

There are four types of *elements* which may be created in a cell. Each element is binary in the sense that each wire or node of an element is capable of two states. (1) A *computing element* may be a switch, a delay, a combined delay and switch, or just a plain wire. Each edge of a cell has at most one computing wire (shown in solid in the figures) impinging on it from within the cell. Specifically, the computing primitives are as follows; see Figure 5. (a) Any switch with three inputs and one output, e.g., cell $A1$ of Figure 5. There are 2^{10} of these, taking into account the different orientations of their outputs (north, south, east, and west), and allowing all 256 possible three-variable truth functions. Note that some of these truth functions will be independent of some of the inputs, so that this category includes as special cases two-input, one-input, and zero-input switches. This category also includes certain wires, such as the wires of cells $A1$, $B1$, and $C1$ of Figure 9. (b) There are unit delay elements with one output and one, two, or three inputs. These inputs come from one, two, or three edges through a disjunction. There are 28 of these; see cell $B1$ of Figure 5. (c) There are unit delays with two outputs and one or two inputs (e.g., cell $C1$ of Figure 5). There are 18 of these. (d) There are unit delays with one input and three outputs ($A2$ of Figure 5); there are four of these since the input can come from one of four different cell edges. (e) There are arrangements of wires in cells not covered by case (a), governed by the

[9] Von Neumann considered the two-dimensional case, but his general procedure applies to any number of dimensions.

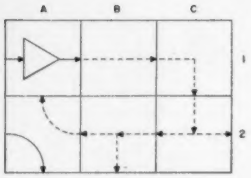


FIG. 6. Scaffolding and constructing primitives.

rule that incoming wires may not merge and at most one wire contacts a side. There are 28 possibilities; see cells $B2$, $C2$, $A3$, $B3$, and $C3$ of Figure 5, and $B4$ of Figure 9. Altogether there are 1102 different computing elements. Systems of these elements may be employed in the usual ways to store and process information.

(2) There are *scaffolding elements*, which have one input wire and either one or two output wires. Scaffolding wires are shown dashed; see cells $B1$, $C1$, $B2$, and $C2$ of Figure 6. Since primitives with different orientations are different there may be any of 24 different scaffolding structures in a cell. A branched scaffolding element (e.g., cell $B2$ of Figure 6) has an implicit switch associated with it. By convention this switch has an initial setting; the switch setting may be changed by using switching sequences which will be introduced later. Scaffolding elements will be used to conduct construction signals from their place of origin to the cell where construction is to take place. (In Sec. 5 scaffolding elements will also be employed to conduct signals which sense the state of a cell.)

(3) Any scaffolding element may coexist with any computing element in the same cell. See, for example, cell $A2$ of Figure 6 and cell $A5$ of Figure 8. There are 24×1102 such primitives.

(4) The last element is the key to the construction process. It is called a *constructing element* and is shown in cell $A1$ of Figure 6. A constructing element receives construction signals from a computing structure and routes these into a scaffolding structure; the construction signals travel though the scaffolding and eventually cause a cell to be structured in a way to be explained. A constructing element may receive its input from any cell edge and transmit its output

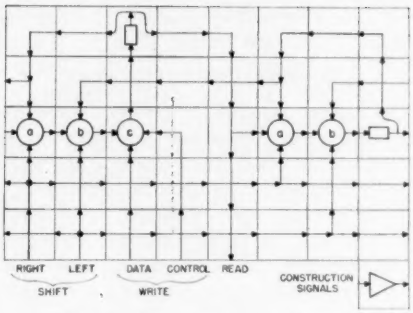


FIG. 7. Two "squares" of tape.

to any other edge of that cell; hence there are 12 different constructing elements.

A cell may have an element of any of these four types or it may be structureless. Altogether there are 27,587 (which is less than 2^{15}) possibilities, each of which will be represented by a binary sequence called a *constructing sequence*. These sequences will be formed or stored in a computer structure and fed into the constructing element. There is no upper limit to the length of computer wires or scaffold wires, and we will make the idealized assumption that signals will travel down the wires instantaneously.

The mode of operation of the system and the formulation of further rules can best be presented by means of an example. We will show how a tape may be lengthened under the control of a machine, thereby relating growing automata to generalized Turing machines. Figure 7 shows two "squares" of the tape, each square storing a single bit. The switching functions (a), (b), (c), of Figure 7 are the following. The output of switch (a) [at the east side of element (a)] will be on if and only if, at the same moment of time, input south is off and input north is on, or input south is on and input west is on. The output of switch (b) [at the east side of (b)] will be on if and only if south is off and west is on, or, south is on and north is on. The output of switch (c) [at the north side of (c)] will be on if and only if east is off and west is on, or east is on and south is on.[10] You are to imagine that a fixed com-

[10] This can be expressed with more precision by employing a symbolic logic notation:
element (a): $E(t)$
$\equiv [\{\overline{S(t)}\ \&\ N(t)\}\ v\ \{S(t)\ \&\ W(t)\}]$

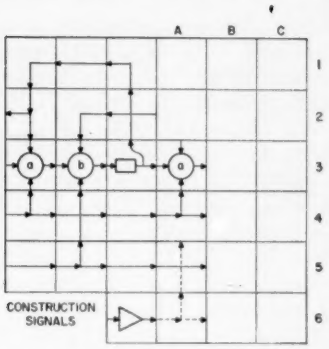


FIG. 8. Construction process.

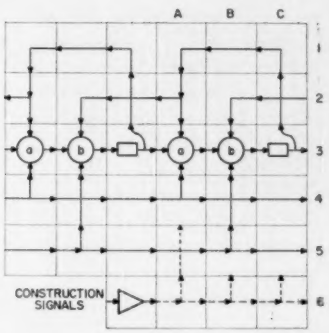


FIG. 9. Completion of new tape "square."

puter which is associated with the tape is connected to the inputs and outputs of Figure 7. This computer may "write" on the left-hand square and "read" what is stored on that square. It may cause the information on the tape to be shifted to the right by stimulating the "right shift" signal, and in a similar way it may cause the information to be shifted to the left. By feeding signals into the constructing element, it can cause the creation, destruction, or alteration of new tape squares to the right in a way to be indicated. A similar arrangement exists for modifying the tape to the left.

We will next describe the construction of a square of tape. Cell $A6$ of Figure 8 needs a scaffold with a branch. The construction sequence representing this element is produced by the fixed computer and fed into the constructing element, which produces the desired element in the cell touched by its output. The next constructing sequence will pass through the constructing element into the scaffold, through the vertical output of the scaffold, and will produce a new element in the cell ($A5$) which is touched by the output of the scaffold. In this case the structure produced is a horizontal computing wire and a vertical scaffolding wire. This process is repeated to give a structure to the next two cells ($A4$ and $A3$); the scaffolding of these cells is not shown because of pictorial difficulties.

After the rest of column A is structured, construction activities need to be shifted to column B of Figure 9. This rerouting is done by a binary *switching sequence* which goes through the constructing element into the scaffold and changes the switch setting of the scaffold of cell $A6$ so that subsequent constructing sequences are routed into cell $B6$. After this we can structure column B by means of six constructing sequences. Another switching sequence resets the switches in cells $A6$ and $A5$ and permits the structuring of column C. This switching sequence works as follows. It would contain two switching signals in coded form. The first signal would be received by $A6$, and in this case would leave the switch set as it was. The second switch signal would be passed on to $B6$, and would cause its switch setting to be changed so that the next constructing sequence would effect cell $C6$ (rather than traveling through the B column). In general, a switching sequence will contain as many switching signals as there are switches which must be passed through to get to the last switch to be modified. (As before, the scaffolding of some cells is not shown.)

The central computer may want to contract the tape as well as to expand it, and so it needs the capacity to wipe out the structure of a cell. This is provided for by binary *destroying sequences*. A destroying sequence passes through the constructing element and down the scaffolding in the same way as a constructing sequence. How-

element (b): $E(t)$
$\equiv [\{\overline{S(t)}\ \&\ W(t)\}\ v\ \{S(t)\ \&\ N(t)\}]$
element (c): $N(t)$
$\equiv [\{\overline{E(t)}\ \&\ W(t)\}\ v\ \{E(t)\ \&\ S(t)\}]$

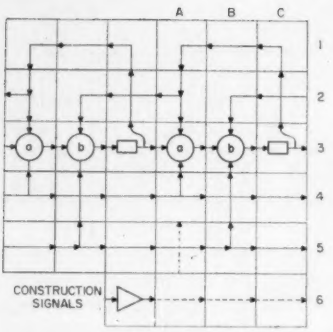


FIG. 10. Preparation for next construction.

ever, a destroying sequence affects the cell which contains the terminus of this scaffolding channel rather than the next cell; it causes the element of this cell to be destroyed. There is a sequence of 18 destroying sequences and 2 switching sequences which will destroy the tape square just created.

We may also use a switching sequence followed by three destroying sequences to erase the elements of the bottom row (6) and then use 3 constructing sequences to create the scaffolding shown in Figure 10. Thus with a master sequence of 21 constructing sequences, 3 switching sequences, and 3 destroying sequences we end up with a new tape square and with the scaffold arranged so that each repetition of this master sequence will produce a new square to the right. By iterating this master sequence we can produce a tape of any length.

In this way we can construct a generalized Turing machine except for the fact that the process of construction just described takes many units of time per tape square and hence is not as fast as that stipulated for a generalized Turing machine in Section 2. Thus the generalized Turing machine is almost but not quite a special case of our definition of automaton. We could make it so in either of two ways. First, we could use a growing automaton to simulate a generalized Turing machine by associating one major cycle (i.e., the time required for the construction of a tape square) of the growing automaton just described with one minor cycle (i.e., a single unit of time) of the generalized Turing machine. Second, we could enrich our base of construction by adding elements and wires capable of more than two states. It would not be necessary to introduce computing elements with more than two states. It would suffice, for example, to let the scaffolding elements have many states and to have constructing elements which convert binary signals into signals properly coded for the scaffolding elements; note that this method would not involve any increase in the number of computing primitives or the number of scaffolding primitives, and not a very great increase in the number of constructing primitives.

We have said enough about the definition of "automata" to make its main features clear. Various details need to be specified before the definition is complete. We will assume that a primitive element is capable of action as soon as it is created. This means that part of a computer which has been constructed may start to operate before the computer is finished. This fact needs to be taken into account in the design of the computer, as it often does in the design of actual computers, since these computers have to be cleared to a standard initial state before they are ready to operate. Two constructing sequences, coming from two different constructing elements, might arrive at a cell at the same time; a priority rule is needed to determine which one governs the construction. There are other minor details which need to be specified but we will not bother with these here.[11]

Let us consider next what automata can be specified as automata in the sense of the present definition. Every fixed automaton can be put into a normal form (Burks and Wang, 1957), and this normal form can be constructed of computing elements in cells;

[11] It is perhaps worth noting that the definition does not prevent the construction of automata which are not well-formed (e.g., have switch cycles), nor is there an effective criterion for deciding whether a given growing automaton will ever "develop" switch cycles. Restrictions could be added to guarantee that every automaton is well-formed. The simplest such restriction would be to incorporate a delay in every switch, but this tremendously complicates the automata and limits their behavior. A less severe restriction is to require every switch or wire with an output on the south edge to contain a delay.

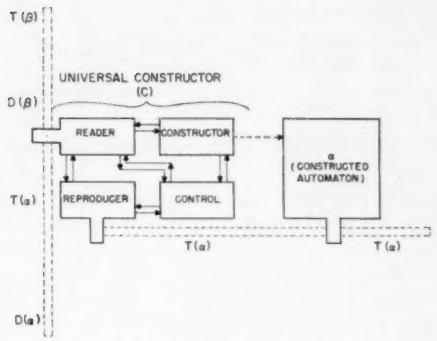


FIG. 11. Universal constructor.

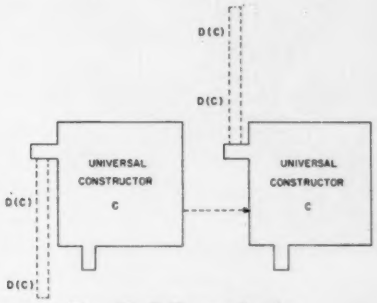


FIG. 12. Self-reproduction.

so each fixed automaton is an automaton in the sense defined here. We have shown in some detail how an expanding and contracting tape may be constructed; hence, except for a matter of speed, every generalized Turing machine is also an automaton.

One of the growing automata which can be constructed is von Neumann's universal constructor (von Neumann, 1958, pp. 2096–2097), of which a block diagram is shown in Figure 11. The *universal constructor* C is a finite structure to which may be attached a tape containing the specifications of any number of automata which are to be synthesized. In general, each specification consists of two parts. The first part of the specification is a master sequence of constructing sequences, switching sequences, and destroying sequences, which will cause a fixed part of the desired automata to be constructed. This master sequence is the description $D(\alpha)$. The second part of the specification contains the contents $T(\alpha)$ of a tape which is to be constructed and attached to automaton α.

The universal constructor operates under the direction of its control as follows. First, the tape reader reads $D(\alpha)$ and transmits it to the constructor; concurrently the constructor constructs α. Next, the reader reads $T(\alpha)$; concurrently the reproducer produces a tape, attaches it to α, and records $T(\alpha)$ on it. This process is then repeated, so that at the end of the construction process $T(\alpha)$ is recorded twice on the tape. (The reason for this duplication of tape information will become apparent in a moment.) In general, for each description $D(\alpha)$ of an automaton and each tape content $T(\alpha)$, the universal constructor will construct α and record $T(\alpha)$, $T(\alpha)$ on the tape attached to α. As mentioned before, α must not be allowed to begin operation before it is finished in such a way that it alters its own construction.

As von Neumann showed, the universal constructor C together with the proper tape will reproduce itself; see Figure 12. Let $D(C)$ be a sequence of constructing sequences, switching sequences, and destroying sequences which completely describe C. This description of C must include whatever scaffolding it has; if C has no scaffolding, then $D(C)$ must be organized in such a way that after a scaffolding has been used it is destroyed. Let the universal constructor C contain a tape storing $D(C)$ twice when it begins action. The universal constructor will first read $D(C)$ and produce C. It will then read $D(C)$ twice and produce a tape recording $D(C)$ each time. [This is the reason for the earlier requirement that $T(\alpha)$ be copied twice.] Thus the universal constructor C when supplied with a tape containing $D(C)$, $D(C)$ produces a universal constructor plus a tape containing $D(C)$, $D(C)$. Hence C with a tape storing $D(C)$, $D(C)$ reproduces itself. The new combination of C and the tape $D(C)$, $D(C)$ can then reproduce itself, and this process can be repeated *ad infinitum*.

If we combine a universal constructor and a universal computer (i.e., a universal generalized Turing machine), and supply this composite machine with a tape containing its description twice, this whole

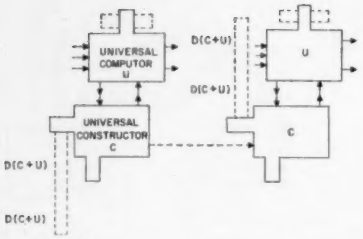


FIG. 13. Universal constructor-and-computer.

automaton will reproduce itself; see Figure 13. Thus a universal constructor-computer can reproduce itself (von Neumann, 1958, p. 2098).

5. SOME PROBLEMS ABOUT GROWING AUTOMATA

In Section 3 we mentioned some problems and results about fixed automata and generalized Turing machines and remarked that these may be extended to cover automata in general (as defined in Section 4) and hence growing automata in particular. There is another type of problem, quite different from any of the problems discussed before, which is applicable to automata of all kinds; this concerns the relation of the cyclic complexity of the structure of an automaton to its behavior. Since our earlier discussion of a generalized Turing machine did not make explicit whatever cycles are implicit in the tapes, this problem had to be postponed to the present section.

Let us consider the problem first for the case of a fixed automaton. We can uniquely decompose the structure of a fixed automaton into *maximal cycles*. The *degree* of a maximal cycle is the number of unit delay elements in it, and the degree of a fixed automaton is the maximum of the degrees of its maximal cycles.[12] Figure 14 consists of a single maximal cycle of degree 2; the output nodes are marked by stars. Figure 15 contains six maximal cycles, each of degree 1, so Figure 15 is of degree 1; its output nodes are also marked by stars.

How, now, is the degree of an automaton related to its behavior? Is there some critical degree n, such that all fixed automata behaviors can be realized by structures of that degree (or less)? It has been conjectured that there is no such degree. The general problem remains open, but John Holland (1959a) has shown that no automaton of degree 1 can realize the behavior of Figure 14, which is a ternary counter.

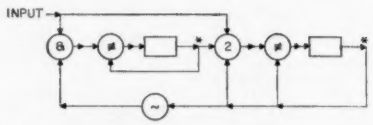


FIG. 14. Automaton of degree 2 (ternary counter).

The same problem can be raised about the computation of a fixed automaton: is there some critical degree m, such that all fixed automata computations can be realized by structures of that degree or less? Both problems can be extended to generalized Turing machines and to growing automata. Of course the structure of any growing automaton changes with time, and hence the degree of the structure changes with time as well. But we can define the degree of a growing automaton to be the maximum degree of any of its structures at any time for any input history, if the maximum exists; otherwise the degree is infinite.

We can now ask these questions about automata generally: (1) Is there some finite degree n such that all behaviors can be realized by automata of this degree or less? (2) Is there some finite degree m such that all computations can be realized by automata of this degree or less? It might seem that the answer to both these questions is "no," since clearly in both cases there is no upper limit to the length of the tapes needed, and the tape designed in the preceding section is of infinite degree. However, it is possible to design a tape of finite degree by using isolated squares and moving the read-write head along the tape instead of shifting the information along the tape. To store a 1 in a square the automaton builds a delay cycle in which a pulse circulates; to store a 0 it builds a delay cycle whose successive states are 0. The fixed part of the machine can send out a long wire to do this and by means of another wire can sense the contents of the delay cycle. The contents of the square

[12] These concepts are explained in detail in Sec. 4.1 of Burks and Wang (1957).

can be changed by destroying the delay cycle in it and rebuilding the appropriate one. When the fixed part of the automaton is neither reading from the square nor writing in the square, it would have no connection to it. In this way a generalized Turing machine may be designed which has a fixed degree. There are other measures of feedback complexity which are of interest. (See Appendix I).

Let us turn next to some problems which are unique to growing automata. The first is one that von Neumann (see Shannon, 1958, p. 127) considered: How complex does a self-reproducing automaton need to be? In his model, von Neumann used cells which were capable of twenty-nine states, and he estimated that about 200,000 cells would suffice for self-reproduction (Kemeny, 1955, p. 66). The cells of our models are much more complicated, for a cell may contain any of about 2^{15} different elements, and each element is capable of several states. But counting states is not sufficient, because there are many ways in which states may be used; for example, a primitive element capable of transferring its structure to a neighboring cell on receipt of a single pulse is in certain respects more powerful than any of the primitives we have assumed.

One could compare growing automata constructed within the framework of different definitions of "automaton" by reducing both definitions to a common system. We will show how to make this reduction for our definition. For this purpose we will use a particular infinite fixed net of switch and delay elements to model or simulate our definition of automaton.[13] This infinite fixed net corresponds to the infinite two-dimensional array of discrete cells constituting the basis of each automaton in the sense of Section 4. You will recall that in Section 4 a particular automaton is defined by specifying the initial structure (i.e., the structure at time zero) of a finite number of cells; this structure would be represented in the infinite fixed net by a particular initial state of that net. In Section 4 the structure of an automaton at time $t + 1$ is determined by the structure at time t and the complete state at t according to rules explained there; this succession of structures (and states) will be represented in the infinite fixed net as a succession of states of that net.

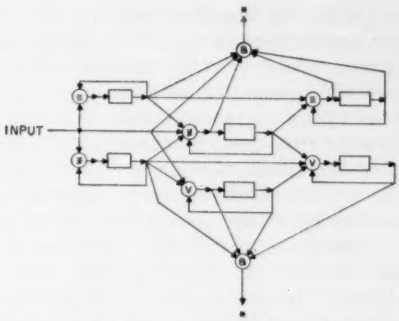


FIG. 15. Automaton of degree 1.

The infinite fixed net is constructed in the following way. There is a finite fixed net (located in a square) corresponding to each cell of our definition of automaton; these squares are repeated *ad infinitum* in the same two-dimensional way as are the cells of Section 4. We thus get an infinite structure which is locally heterogeneous but is homogeneous in the large.

We will indicate only the main features of the finite fixed automaton corresponding to a cell. It will contain a register, a master switch, and switches and delays which will perform the scaffolding and computer functions of the original cell. Each fixed net of a square would have two inputs to, and two outputs from, each of the four neighboring squares; these inputs and outputs would be used for scaffolding wires and computer wires. A constructing sequence which was to affect a given square would be routed to the register of that square; the register would then determine the connections of the switches and delays by means of the master switch; in this way the fixed automaton of the square would simulate the original cell. Flip-flops would be used to store the switch settings of branched scaffolding elements; a switching sequence passing through a square would set the flip-flop correctly and then

[13] Note that though the net is infinite, at any finite time only a finite number of nodes have ever changed from their initial state; cf. the remark on infinite tapes in Section 2, note 4. For other definitions of growing automata see Church (1958), and Holland (1959b).

pass on to set the flip-flops representing other scaffolding branches. A destroying sequence would go into the register and restore the square to its original state.

Each particular automaton is defined within the framework of our original definition by stipulating the initial structure of a finite number of cells. As we remarked before, this operation corresponds to setting the initial states of the delay elements of a finite number of squares in the infinite fixed net. Changes of automaton structure in the original definition will then appear as changes of state in the infinite fixed net. Thus we are using some aspects of the behavior of the infinite fixed net to simulate the structure of the original automaton.

Von Neumann's definition of automaton could be simulated in a similar way. Then the complexity of von Neumann's design of a self-reproducing automaton could be compared with the complexity of the design we described by making a weighted count of the number of switches and delays used in the initial structure of each. By precisely defining a common reference system in this way, we can make the problem of minimal complexity of cell reproduction into a strictly logical problem. Experience with minimality problems suggests that it would be extremely difficult if not impossible to find the minimal solution and prove it minimal. One might, however, obtain an estimate of minimal complexity by using a computer to help work out many alternative constructions.

The idea of modeling a specific structure in an infinite uniform fixed net may have practical applications in future computers. There are under development radically new techniques for constructing solid-state computers. It is estimated that by one of these methods "the component density in a finished circuit could be 50 million per square inch per layer. . . . It is conceivable that more than 10,000 layers per inch could be formed, giving a volume density of 5×10^{11} components per cubic inch" (Buck & Shoulders, 1959, p. 2). Such computers might be built for the most part in a uniform way similar to that described above for the infinite fixed net used to simulate our definition of automaton. A particular problem could be solved by converting this net into a "structure" appropriate to that problem, in the way the infinite fixed net is structured to represent a particular growing automaton. For each new problem (or portion of a problem) the computer would simulate a different structure; in effect, it would in a sense become (during the solution of this problem) a special purpose computer especially designed to solve that problem. The common division of a computer into arithmetic units, controls, and storage units would be used only if this organization was best for the problem at hand. All this simulated restructuring would be done automatically, of course, by a programmed change of state of the uniform fixed net, and could be done more or less continuously during the solution of a problem.[14]

There are many other concepts of growing automata worthy of investigation, of which we will mention a few. It is not necessary to employ a fixed framework of cells; one could allow new cells to spring up between old ones under the control of the automaton. Suppose a list of words is stored and at a later date many words are to be lengthened and new entries are to be inserted. This change could be conceived as an automatic process of lengthening the bins which hold the word, and the inserting of new bins between the old ones; the change of bins must of course be accompanied by an appropriate change of the switches which connect these bins to the rest of the automaton. In general, storage and computing facilities would be created wherever needed and in a form suited to the problem being solved.

In one of his models, von Neumann had girders, sensing organs, and joining organs, as well as computing elements, floating freely in a lake. As Shannon (1958, pp. 126–127) remarks, because of the complexity of motion which is thereby possible it would

[14] Similarly, we can think of the program of a present-day general purpose computer as determining a temporary structure, and hence as converting the general purpose machine into a special purpose machine for the given problem. But the computer described above would have much more flexibility in this respect than any existing computer.

be exceedingly difficult to give a detailed design of a universal constructor in this model. One might modify the definition of automaton we have given (Section 4) in at least two ways to obtain some of the effects of von Neumann's "lake" model without all of the complexity. In the first case the result is a probabilistic automaton, while in the second case it is a deterministic automaton.

(1) We can assume as primitives a few basic types of cells, each with a number of internal states. These cells would be moving about on the surface of a lake in a random fashion. The random motion would cause cells to make contact with one another along the edges. Each cell would have the capacity to accept or reject other cells under the control of its own internal state. A particular automaton, such as a self-reproducing one, would begin life as a two-dimensional finite structure floating on the surface. The cells on its edges would contact other cells at random, and the automaton would accept or reject these new cells according to its needs. Stimulated by von Neumann's work, L. S. Penrose (1959a, b) has investigated models of this sort, using cells with mechanical hooks and producing random motions by shaking the cells on a board. His units are capable of only a few internal states, and the devices he begins with consist of only a few units.[15] Penrose did his work experimentally, that is, by constructing the units from wood and shaking them mechanically. This general type of model could be simulated on a digital computer.[16]

It would be instructive to make such simulations, starting with cells, rules for interconnections, and initially given automata, all of which are much more complicated than the units, rules, and initial structures Penrose used.

[15] Most of his models were one-dimensional, although he did some work on a two-dimensional model.

[16] The digital computer is (when working properly) a deterministic automaton, while the model simulated is probabilistic. The required randomness could be achieved by using a table of random numbers. The simulation will still be deterministic, but it will sufficiently approximate the probabilistic model to be satisfactory.

(2) Von Neumann's "lake model" and the variations of it just discussed are probabilistic automata. A second way of modifying our definition of automaton (Section 4) so as to obtain some of the power of von Neumann's lake model without all the complexity yields a deterministic automaton. The modification consists in adding a *sensing element* which would sense the structure and contents (state) of a cell. This sensing element could use the scaffolding described in Section 4 to send out sensing signals to the cells at the end of the scaffolding; these signals would later return by the same route with information about the cells sensed.

By a process analogous to the construction procedure described in the previous section, an automaton could sense what element was contained in a cell and the state of that element at that time. This information could be used in various ways, e.g., it might be used to detect a fault which could later be repaired by the constructing element.

The question then arises: How much self-knowledge can an automaton acquire and keep? This question is of interest to the theory of self-organizing systems since a possible procedure in self-organization is for the system to have a picture of its structure and behavior and use this picture to guide itself. The automaton might be given this picture initially or it might be programmed to construct the picture by means of sensing elements. Clearly there are limits to how much knowledge an automaton can acquire and keep about itself. For example, a growing automaton might change faster than its sensing elements could sense these changes. There are also logical limitations. Thus if all the states of an automaton are used, a proper part of the system cannot have a completely detailed picture of the whole, for the whole, being larger than the proper part, would be capable of more states than this part.

The logical paradoxes of self-reference are relevant here. It is well known that some referential statements lead to contradictions. Thus the sentence, "This sentence is false," is false if true (because it says of itself that it is false) and true if false (be-

cause if it's false that it's false, then it is true), and hence both true and false. To resolve such paradoxes, Russell (Russell & Whitehead, 1925, Ch. II of the Introduction, "Theory of Logical Types") proposed a formal language in which no self-referential statements could be formulated. The proposal that no self-referential statements be allowed in a formal system is much too strong, however, for there is clearly nothing wrong with the self-referential statements: (1) "This sentence is in English"; (2) "This sentence is written at least once." Indeed, both of these sentences are true. These sentences illustrate the fact that a sentence can refer to its physical properties without contradiction. There are also cases where a sentence can refer to its structural or syntactical properties without contradiction. The sentence, "This sentence is grammatical," is a case in point. Another case of a formula consistently referring to its own structure is the undecidable formula constructed by Gödel (1931) in his proof of the incompleteness of arithmetic. This is a formula F which says, in a certain sense, "F is not provable." (See Appendix II.) Provability is a purely syntactical concept so this formula is referring to its own syntax, and it does so without contradiction.

Thus a sentence may contain a description of some of its own syntactical or structural properties. Note the resemblance of this to what we found in the preceding section, namely, that a self-reproducing automaton C contains within itself a description $D(C)$ of its own structure. It does not, however, contain a description of its own state or contents (except for its initial state). This suggests that an automaton with a sensing element might be able to sense its own structure and construct and store a description of its structure. There is a problem in designing an automaton to do this, however.

Prima facie it might seem that an automaton could not store a description of its own structure because, however many cells it had, storage of the description would require more than that number of cells, in the manner of the Tristram Shandy paradox.[17] This objection is of course not sound, because we may use indices, summation signs, and quantifiers in the description. You will recall that a certain master sequence is used to describe one square of the tape. In writing the description $D(\alpha)$ of the structure of automaton α one does not repeat this master sequence for each square of tape to be constructed, but rather uses the same sequence repeatedly in a manner familiar to all programmers. This can be done because the tape has a highly uniform structure.

It follows from these considerations that in order to produce a description of its own structure which is sufficiently small to be stored in that structure an automaton must detect some of the uniformities of its structure. It could, for example, systematically sense the structure of each of its cells and then analyze the results in order to find a more compact statement of them. Hence the amount of information a growing automaton can acquire and store about its own structure depends on what mechanical techniques there are for discovering uniformities of structure.

[17] Tristram Shandy took two years to write the story of the first two days of his life. (Sterne, 1928, Vol. 4, Ch. 13.)

Appendix I. A Measure of Feedback Complexity

Consider a fixed automaton with n unit delays $x_1, x_2, x_3, \ldots, x_n$. Let $f(x_i)$ be the number of delays which directly and immediately effect the delay element x_i, i.e., whose outputs drive the input of x_i through switch elements. [In Figure 14, $f(x_1) = f(x_2) = 2$ since each delay drives itself and the other delay through a switch.] Define an average measure of feedback complexity F for the automaton by

$$F = \frac{\sum_{i=1}^{n} f(x_i)}{n^2}$$

There are "directly connected" automata for which $F = 1$ (Burks & Wang, 1957, p. 291).* Figure 14 shows such an automaton. However, the amount of feedback complexity F in actual systems is usually much less than the maximum. $F = 7/18$ for Figure 15 and F is small in a typical modern digital computer. (Note that such a computer is pretty nearly one maximal cycle,

* Note: Of course every automaton can be put in a normal form which has this property, but we are here interested in nets in "reduced form," in which all irrelevant (noneffective) switch input wires are deleted.

not counting the tapes.) This fact is relevant to our ability to understand a large digital computer. Because the ratio of actual feedback lines (from delays) to possible feedback lines is low the various parts of the computer have a large amount of local autonomy and may be understood in isolation from the rest of the computer. (Uniformities within these parts also make it easier for us to understand them.) It is worth noting that if neurons are represented by switches and delays, F is small for the human neural system.

Appendix II. Problems of Self-reference

In his lectures at the Institute for Advanced Study in 1934 (as reported in some mimeographed notes entitled "On Undecidable Propositions of Formal Mathematical Systems"), Gödel pointed out that his undecidable formula showed that Russell's solution of the paradoxes was too extreme because the undecidable formula (truly) says of itself that it is not provable. He agreed with Russell that there must be some limitation on what a sentence may say about itself and suggested that no sentence can properly talk about its own truth and falsity.

This idea may be generalized. The relation between a symbol (whether a word, phrase, or sentence) and what it refers to or designates is sometimes called the name relation. It is construed sufficiently broadly to include the relation between a sentence and its truth value and the relation between a predicate and the things it may be predicated of, as well as the relation between a proper name and what it names. Thus Gödel's suggestion that no sentence can properly talk about its own truth and falsity is a special case of the following more general principle: a symbol cannot refer to the relation between itself and what it refers to or designates. To put the point in another way: a symbol cannot properly refer to its own meaning. On the other hand, the examples of consistent self-reference given in the text show that a symbol may properly refer to itself or its physical properties and that it may (in suitable circumstances) refer to its syntactical (structural) properties or to its syntactical relations to other symbols. To summarize: semantic self-reference is to be avoided, but certain syntactical and physical self-references are allowable. (See Reach, 1938).

REFERENCES

Büchi, J. R., & Elgot, C. C. Decision problems of weak second order arithmetics and finite automata, Part I. *Notices Am. Math. Soc.*, December 1958, 5, 834. (Abstract)

Büchi, J. R., Elgot, C. C., & Wright, J. B. The non-existence of certain algorithms of finite automata theory. *Notices Am. Math. Soc.*, February 1958, 5, 98. (Abstract)

Buck, D., & Shoulders, K. An approach to microminiature printed systems. Notes from summer session course in Advance Theory of Logical Design of Digital Computers, The University of Michigan, 1959. This paper was presented at the Eastern Joint Computer Conference, Philadelphia, Pa., 1958.

Burks, A. W. The logic of fixed and growing automata. In H. Aiken (Ed.), *Proceedings of the symposium on the theory of switching*. Cambridge, Mass.: Harvard Univ. Press, 1959, 147–188.

Burks, A. W., & Wang, H. The logic of automata. *J. Assoc. Computing Machinery*, 1957, 4, 193–218 and 279–297.

Burks, A. W., & Wright, J. B. Theory of logical nets. *Proc. IRE*, 1953, 41, 1357–1365.

Burks, A. W., & Wright, J. B. Unpublished research results.

Church, A. Application of recursive arithmetic in the theory of computers and automata. Notes from summer session course in Advance Theory of the Logical Design of Digital Computers, The University of Michigan, 1958.

Copi, I. M., Elgot, C. C., & Wright, J. B. Realization of events by logical nets. *J. Assoc. Computing Machinery*, 1958, 5, 181–196.

Davis, M. *Computability and unsolvability*. New York, N. Y.: McGraw-Hill, 1958.

Elgot, C. C. Decision problems of weak second order arithmetics and finite automata, Part II. *Notices Am. Math. Soc.*, February 1959, 6, 48. (Abstract)

Gödel, K. Uber Formal Unentscheidbare Satze der *Principia Mathematica* und Verwandter Systeme. *Monatsh. Math. u. Physik.*, 1931, 38, 173–198.

Holland, J. H. *Cycles in logical nets*. Ph.D. dissertation, The University of Michigan, Ann Arbor, Mich., 1959. (a)

Holland, J. H. A universal computer capable of executing an arbitrary number of subprograms simultaneously. *Proc. Eastern Joint Computer Conference*, 1959. (b)

Kemeny, J. G. Man viewed as a machine. *Sci. American*, April 1955, 192, 58–67.

Kleene, S. C. *Introduction to metamathematics*. New York, N. Y.: Van Nostrand, 1952.

Moore, E. F. Gedanken-experiments on sequential machines. In C. E. Shannon and J. McCarthy (Eds.), *Automata studies*. Princeton, N. J.: Princeton Univ. Press, 1956, 129–153.

Penrose, L. S. Automatic mechanical self-reproduction. In *New Biology*, Penguin Books, 1959, 92–117. (a)

Penrose, L. S. Self-reproducing machines. *Sci. American*, June 1959, 200, 105–114. (b)

Reach, K. The name relation and the logical antinomies. *J. Symbolic Logic*, 1938, 3, 97–111.

Russell, B., & Whitehead, A. N. *Principia mathe-*

matica. (2nd ed.) Cambridge, England: Cambridge Univ. Press, 1925.

Shannon, C. E. Von Neumann's contributions to automata theory. *Bull. Am. Math. Soc.*, May 1958, 64, 123–129.

Sterne, L. *The life and opinions of Tristram Shandy, Gentleman*. New York, N. Y.: Random House, Modern Library, 1928.

Turing, A. M. On computable numbers, with an application to the Entscheidungs problem. *Proceedings of the London Mathematical Society*, Series 2, 1936, 42, 230–265; 1937, 43, 544–546.

von Neumann, J. The general and logical theory of automata. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior—The Hixon Symposium*. New York: Wiley, 1951. Reprinted in James R. Newman (Ed.), *The world of mathematics*. New York: Simon and Schuster, 1958, 4, 2070–2098.

von Neumann, J. Probabilistic logics and the synthesis of reliable organisms from unreliable components. In C. E. Shannon and J. McCarthy (Eds.), *Automata studies*. Princeton, N. J.: Princeton University Press, 1956, 43–98.

Wright, J. B. Unpublished research results.

(Manuscript received November 21, 1959.)

Man finds himself living in an aleatory world; his existence involves, to put it badly, a gamble. The world is a scene of risk; it is uncertain, unstable, uncannily unstable. Its dangers are irregular, inconsistent, not to be counted upon as to their times and seasons. . . . Our magical safeguard against the uncertain character of the world is to deny the existence of chance, to mumble universal and necessary law, the ubiquity of cause and effect, the uniformity of nature, universal progress, and the inherent rationality of the universe. Those magic formulae borrow their potency from conditions that are not magical. Through science we have secured a degree of power of prediction and control. . . . But when all is said and done, the fundamentally hazardous character of the world is not seriously modified, much less eliminated.

JOHN DEWEY, *Experience and Nature*

In making supposedly rational decisions, we must estimate at least two types of quantity, namely, the relative probabilities of different outcomes of alternative courses of action and the degree of our preference for the different outcomes. The extent to which the fear of failure deters us and the hope of success spurs us on, as well as the degree of our confidence in what may happen, all determine our decisions. In this article, the values of parameters related to the factors influencing judgment are compared in young and old people.

ASPECTS OF JUDGMENT AND DECISION MAKING: INTERRELATIONSHIPS AND CHANGES WITH AGE

by Michael A. Wallach and Nathan Kogan[1]

Massachusetts Institute of Technology and The Educational Testing Service

THE present paper has two major concerns: (1) the investigation of empirical relationships among variables derived from the domains of decision making and the psychology of judgment; (2) the examination of age differences in these domains.

Relationships Among Aspects of Judgment and Decision Making

Extremity and confidence of judgment. Recently, a number of investigators have speculated about the possibility that individual differences in the judgmental realm may reflect different ways of minimizing error. Thus, Pettigrew (1958), Wallach and Caron (1959), and Bruner and Tajfel (1960) suggest that categorizing (judgment) style may involve such considerations as an individual's tolerance limits for error. In a similar vein, Cohen (1958) remarks that "a degree of doubt is manifested not only when we make decisions but wherever we express a preference, make an estimate, extrapolate into the future or interpret the past" (p. 286). A first empirical attempt (Kogan and Wallach, 1960) to relate a judgment measure —confidence—to an independent, albeit phenomenological, measure of risk showed that individuals with lower confidence scores tended to evaluate such concepts as "risk" and "stock market" as more hostile, cold, and tense on a semantic differential (Osgood, Suci, & Tannenbaum, 1957). Finally, note should be taken of those studies (Gardner, 1953; Hamilton, 1957) that claim to have found a relationship between preference for narrow inclusion categories in perceptual judgments and a tendency toward obsessive, anxious doubting.

In the present paper, we shall be concerned with (1) the extremity aspect of judgment, i.e., an individual's predisposition toward categories located near the extremes of a distribution of alternative categories in a given situation; and (2) judgment confidence,[2] i.e., an individual's degree of conviction that the judgment he has made or is

[1] This investigation was supported by a research grant (M-2269) from the National Institute of Mental Health, Public Health Service, conducted under the auspices of The Age Center of New England, Inc. Grateful thanks are due Marguerite Braun, Ruthellen C. Gahm, Leonard R. Green, and Lynne Hamilton for aid in gathering and analysis of data; Arthur S. Couch for aid with machine computation methods; Stephen R. Deane, Donald S. Dunbar, Bernhardt Lieberman, A. Bertrand Warren, and Harold S. Zamansky for making classes available; and Samuel J. Messick for critical comments. Data for this study were processed at the MIT Computation Center.

[2] In our previous work (Wallach and Kogan, 1959; Kogan and Wallach, 1960), we used the term "certainty." Since that term has a highly specific meaning in decision theory, we have here adopted the term "confidence" in the interest of avoiding conceptual confusion.

about to make is the most appropriate one under the circumstances. Prior evidence indicates that these two variables are positively correlated (Brim & Hoff, 1957; Wallach & Kogan, 1959). The present work will permit a replication on an older sample.

Utility and subjective probability in decision making. The studies cited make it clear that judgmental processes may well have a motivational basis in terms of (1) perceived positive and negative consequences of judgments in particular contexts, and (2) a generalized consistently cautious or risky outlook affecting an individual's judgments in a variety of situations. When formulated in this manner, concepts derived from the psychology of judgment (and cognitive processes, more generally) may lend themselves to articulation with concepts associated with decision-making theory. In regard to the latter, we are particularly interested in the concepts of utility and subjective probability in decision making under uncertainty (Luce & Raiffa, 1957).

Feather (1959) reviews five psychological approaches to this problem and notes that decision models based on chance-related situations assume independence between utility and subjective probability, while models of choice behavior in level of aspiration and achievement contexts assume a relationship. It should be noted that the above approaches are concerned with the role of these two concepts in affecting the choices made in particular decision situations. In the present study, we shall pursue an individual differences approach and inquire whether utility and subjective probability as concepts descriptive of individuals bear any relationship to each other. Techniques were used permitting an independent measurement of the extent to which individuals considered success desirable (relative to the disutility of failure) and success more likely (relative to the probability of failure).

Further relationships. Here we consider possible relations between extremity and confidence of judgment on the one hand, and utility and subjective probability of success and failure on the other. More specifically, we shall test the hypothesis that for individuals disposed toward greater extremity and confidence in their judgments, success will have greater utility (relative to the deterrence or disutility of failure) and/or a greater subjective probability of attainment.

Age Differences in Judgment and Decision Making

Extremity and confidence of judgment. Since prior research indicates that extremity and confidence of judgment are positively correlated, determinants of differences in judgment extremity other than confidence level itself can be assessed only if confidence is held constant. Assessment of the relationship between extremity and age hence requires the control of confidence level. We shall also inquire into young-old differences in the over-all confidence with which judgments are made.

Such an examination of extremity-confidence in relation to age might help to clarify some confusing results obtained by Korchin and Basowitz (1956) in an "intolerance for ambiguity" experiment. In that study, *Ss* were required to respond "dog" or "cat" to a graded series of pictorial stimuli varying in the degree to which one or the other animal quality was present. The authors hypothesized a greater intolerance for ambiguity on the part of the older individual, with the result that he would tend to "seek black-white alternatives." Their hypothesis, in other words, states that older individuals will make more extreme responses in ambiguous situations than younger persons. The results of the study do not seem to confirm the hypothesis, for older *Ss* exhibited considerably more vacillation in dealing with the most ambiguous photographs. While the younger *S* shifted abruptly from "cat" to "dog" somewhere in the middle of the stimulus series—i.e., switched from one extreme to the other—the older person seemed to have a more difficult time deciding whether these ambiguous stimuli were "cats" or "dogs." In other words, older *Ss* were less extreme in their judgments.

Perhaps the major difficulty here is that confidence and extremity in the Korchin-

Basowitz study cannot be empirically separated. Old and young individuals may be equally confident in their judgments, with the younger more prone to go to extremes under these conditions; or the young *Ss* may manifest more extreme responses for the very reason that they are more confident. The present investigation should assist in resolving this problem by examining age differences in extremity of judgment within confidence levels rendered subjectively equal in magnitude for young and old *Ss* of the same sex.

Utility of success vs. disutility of failure. Experimental comparisons of young and old individuals in cognitive functioning suggest that aging is responsible for significant decrements in performance. Such decrements have been reported, for example, in paired-associate learning tasks (Ruch, 1934; Korchin & Basowitz, 1957) and in the capacity to arrive at and to break perceptual closure (Basowitz & Korchin, 1957). In both of these areas, the inferior performance of the older individual can be traced, in large part, to response inhibition, i.e., an apparent preference for errors of omission rather than errors of commission. On the basis of this observation, Basowitz and Korchin (1957) infer that the aged individual is characterized by greater caution.

In dealing with learning and perceptual phenomena, however, an inferred mediating construct of caution in decisions must compete with other explanations. For example, age decrements in the cognitive performances described may reflect the inadequate time allotted the older person, given the slowing of sensori-motor functions typical of old age (Birren, 1955). Or, as Botwinick (1959) has indicated, diminished sensory and perceptual capacity may be responsible for reducing the number of cues extracted from the test material. But if it can be shown that older individuals score higher than younger on an independent measure of the deterrence or disutility of failure (relative to the utility of success), then caution as a mediating factor can more readily be invoked in accounting for age differences in other cognitive areas. It would be demonstrated that at least one aspect of caution—namely, the disutility of failure—increases with age.

Does one find a general increase in the disutility of failure over the utility of success for older individuals as compared with young adults, and is this true both for males and females? In addition to such a general increase, or as an alternative to it, does one find that changes in utility with age for each sex depend on the nature of the particular failure and success in question?

Age comparisons within sex. Since our earlier work (Wallach & Kogan, 1959) found differences between young adult men and women in the aspects of judgment and decision making discussed above, the impact of the aging process on confidence, extremity, and the disutility of failure will be examined separately for each sex.[3]

METHOD

Samples

A total of 511 persons comprising four subgroups participated in the present experimentation: 89 older women; 132 younger women; 65 older men; and 225 younger men. All procedures were administered individually to the older persons and in groups to the young adults. With regard to age of the older *Ss*, the mean age for older men was 70.2 ($S.D. = 7.3$), while the mean age for older women was 69.5 ($S.D. = 7.7$). The young adult samples were undergraduates enrolled in introductory psychology courses. The males were from Northeastern University, and the females from Simmons College. Both schools are in the Boston area, and comparable with regard to the socioeconomic status and intelligence of their students. Our intention for purposes of age comparison work was to obtain samples of older individuals who would be as closely matched as possible to the young adults. It was obvious, therefore, that samples of older persons biased toward the upper end of the education and intelligence scales were required. The *Ss* available at The Age Center of New England, Inc., a non-profit gerontological research organiza-

[3] Data on subjective probability estimates were unavailable for our young adult subjects, thus precluding an examination of age differences on this variable.

tion in Boston, seemed to fulfill this criterion quite well.

The degree of intellectual match achieved across age groups may be evaluated from the following data. The mean highest attained education level for older men was 13.7 years of schooling (*S.D.* = 2.8), while the mean for older women was 12.9 years of schooling (*S.D.* = 3.0). Since 12 represented the senior year of high school and 16 the senior year of college, this meant that the older men had attained a mean of 1.7 years of college, while the older women had reached a mean of 0.9 years of college. This placed the older samples at about the same average number of college years as had been achieved by our younger groups, since the young adult samples consisted mostly of college freshmen, with a few college sophomores.

In addition, an intelligence measure was obtained from the older samples: the vocabulary subtest of the Wechsler intelligence test for adults (Wechsler, 1944). This vocabulary subtest has a very high correlation with the combined score from the complete Wechsler battery. Our major objective in administering the vocabulary measure here was to determine the intellectual comparability of the older *Ss* to our young adult samples. Since Doppelt and Wallace (1955), Strother, Schaie and Horst (1957), and others, have found that vocabulary scores show the least decline with age of all intelligence measures, the Wechsler vocabulary subtest seemed most appropriate for comparing intelligence across ages.

Wechsler's manual of instructions was followed in the administration and scoring of the vocabulary test. Two independent judges scored all vocabulary tests. Any scoring disagreement between the two judges' sum vocabulary credit scores for an *S* was resolved by taking their mean. Disagreements were minimal, however, as witnessed by an intraclass correlation coefficient between the two judges' sum scores for all older *Ss* of .95. As Robinson (1957; 1959) has shown, the intraclass correlation coefficient is a highly appropriate measure of reliability, since it measures degree of agreement in absolute level as well as degree of correlation between two sets of scores.

The mean vocabulary score for older men was 34.7 (*S.D.* = 4.7), while the mean for women was 33.9 (*S.D.* = 5.2). Data from Dana (1957) and others on college students indicate that the scores of these older *Ss*, if anything, exceed those of young adult college students.

The older *Ss* were in apparently good physical health, not institutionalized, and were quite capable of coping with the demands of the present experimental task. Relative to college students, the present sample of older *Ss* probably represents as similar a comparison group as one can obtain in practice.

Extremity Index

Extremity of judgment was assessed by a technique discussed in detail in an earlier report (Wallach & Kogan, 1959), and derived from Brim (1955). In brief, a questionnaire was given to *S* requiring judgments about the likelihood of various events. The actual range of content of the items was quite broad, and most items were so ambiguous that no correct answer could be specified. Each of the 50 items in this questionnaire was of the form: "The chances that such-and-such event will occur are about ____ in 100." The *S*'s task was to give an opinion as to the likelihood of the event in question by estimating the number of chances out of 100 that the event will occur. Given the general ambiguity of the items, as probability estimates move away from zero or 100 chances out of 100, *S*'s judgments become manifestly less extreme. In order to measure extremity of estimate regardless of whether its direction was toward zero or 100, each of *S*'s estimates was subtracted from zero if the estimate was below 50, and from 100 if the estimate was above 50. Extremity scores thus could range from zero (maximum extremity) to 50 (minimum extremity).

To control for level of confidence, *S*, after making each judgment, had to indicate how sure he was of this judgment by circling one of the following five confidence categories listed beneath the item:

very sure	quite sure	moderately sure	slightly sure	not sure at all

An average extremity score was calculated for judgments placed at each level of confidence, yielding five extremity scores for each *S*.

Confidence Index

The confidence categories listed above were weighted from one to five, respectively. The number of judgments marked at each confidence level was multiplied by the appropriate weight and the products summed across the five confidence categories, yielding a single index of confidence for each individual. The smaller this index, the higher *S*'s over-all level of confidence.

Deterrence of Failure Index

A questionnaire index of disutility or deterrence of failure was constructed (Wallach & Kogan, 1959). In the questionnaire, twelve everyday life situations are described to *S*. The central person in each situation is faced with a choice between two courses of action, one of which is more risky than the other but also is more rewarding if successful. The *S* must indicate how desirable he feels it would be for the person described to follow the more risky course of action and to succeed, by selecting the probability of success that is deemed sufficient to warrant choosing the risky alternative. Hence this index measures the relative disutility or deterrence of failure vs. the utility of success regarding the more risky alternative.

As an example of the situations presented, the first item follows in its entirety:

"(1) Mr. A, an electrical engineer who is married and has one child, has been working for a large electronics corporation since graduating from college five years ago. He is assured of a lifetime job with a modest, though adequate, salary, and liberal pension benefits upon retirement. On the other hand, it is very unlikely that his salary will increase much before he retires. While attending a convention, Mr. A is offered a job with a small, newly founded company with a highly uncertain future. The new job would pay more to start and would offer the possibility of a share in the ownership if the company survived the competition of the larger firms.

Imagine that you are advising Mr. A. Listed below are several probabilities or odds of the new company's proving financially sound. *Please check the lowest probability that you would consider acceptable to make it worthwhile for Mr. A. to take the new job.*

____The chances are 1 in 10 that the company will prove financially sound.
____The chances are 3 in 10 that the company will prove financially sound.
____The chances are 5 in 10 that the company will prove financially sound.
____The chances are 7 in 10 that the company will prove financially sound.
____The chances are 9 in 10 that the company will prove financially sound.
____Place a check here if you think Mr. A should *not* take the new job no matter what the probabilities."

In brief, the eleven remaining items concerned the following situations: (2) a man with a heart ailment who must choose between changing his life habits or undergoing a dangerous medical operation; (3) a man who may invest in secure, low-return stocks or in risky securities which offer the chance for large gains; (4) a football captain who can choose a safe play which will tie the score or a more risky play which might bring victory; (5) a corporation president who may build additional facilities in America with the result of a moderate return on the investment, or build additional facilities in a foreign country with an unstable political history, but where returns would be much higher; (6) a college senior with the choice of attending a graduate school with very high standing where he might flunk out, or a graduate school with a lesser reputation where he would be sure to pass; (7) a chess player in a national tournament who has the chance to employ a risky strategy which would bring victory if successful, but quick defeat if not; (8) a young man of considerable musical talent who must decide whether he wishes to enter medical school, or a conservatory of music for further training in a field of uncertain success; (9) an American prisoner of war who has the chance to escape, with the risk of execution if caught; (10) a man who must decide whether or not to run for Congress in a hot campaign filled with attacks by his opponents; (11) a research physicist with the choice of working on a difficult but very important problem with the risk of complete

failure, or on easy but much less important problems where success would be assured; (12) a couple who must decide if they should marry despite recent indications of sharp differences of opinion.

Each of these eleven situations was presented to *S* in about the same degree of detail as the first item. The response categories were arrayed from chances of one in ten upward for the odd items and from high probabilities down to chances of one in ten for the even items, thus counterbalancing for any order preferences in choice of probability levels. Refusal to recommend the risky alternative no matter what its probability of success was scored as ten. The larger the probability level selected, the greater the deterrence value of failure for *S*. Individual item scores were obtained, and, in addition, an average probability level for the twelve items was computed (hereafter called the over-all deterrence of failure score).

Subjective Probability of Failure Index[4]

Subjective probability scores were derived from the following procedure: *S* is shown a toy wind-up car placed at a point marked "X" on a table. Point "X" forms an isosceles triangle with two toy posts marking the corners of the triangle's base. The distance from the center of this base to point "X" is three feet, and the two toy posts forming the corners of the base are 12 inches apart. After *E* demonstrates that the car runs in a straight line when released, *S* is told that his task will be to aim the car from the point marked "X" in such a way that it goes between the two posts without touching either of them. The instructions then continue:

"You'll have a chance to try in a few minutes; but first we want to get your judgments about how well you *think* you'll be able to do. For instance, with the posts set this distance apart, how many times out of five tries do you think you could aim the car so that it would go between the posts without touching either of them?"

[4] The impetus which led to the development of this procedure came from a study by Cohen, Dearnaley, and Hansel (1956) on the training of men to drive double-decker buses in England.

For this question, the posts were set at their maximum separation distance of 12 inches. After *S* makes his estimate, *E* brings each post half an inch closer, and asks: "Now suppose we set the posts like this. How many times out of five this time?"

After *S*'s estimate, each post is again brought half an inch closer and *E* repeats the question. In this manner, the posts are gradually brought together an inch at a time, with a further prediction of the number of successes out of five tries being obtained at each smaller separation until a point is reached where *S* replies that he could never get the car to clear the posts in five tries.

Then *E* moves the posts apart again to their maximum separation distance of 12 inches. He asks *S* the following:

"Now this time, instead of giving me a number, I just want you to tell me whether you feel *sure* or *not sure* that, on the first shot, you could get the car between the posts without touching either post. Are you sure, with the posts set like this?"

If *S* decides "sure," each post is brought half an inch closer and *E* asks, "Are you sure here?" If *S* again decides "sure," each post again is brought half an inch closer and the question repeated. This sequence is continued until *S* finally decides "not sure."

A performance series now follows in which *S* actually tries to aim the car. The posts are first set their maximal distance apart, and *S* has a try. *E* then moves each post half an inch closer and *S* has another try. *E* continues to move each post half an inch closer after each successive try until the car strikes one or the other post on three successive tries. Then the posts are moved apart once more to their maximum separation distance and this performance series again is repeated until the car once more strikes either post three times in a row.

The following estimation measures were derived from the subjective probability of failure procedure: (1) The number of inches separating the posts when *S last* predicted that he would succeed five times out of five tries. (2) The number of inches separating the posts when *S first* predicted that he would succeed no times out of five tries. (3) The number of inches separating the

TABLE 1
JUDGMENT EXTREMITY AT EACH CONFIDENCE LEVEL FOR YOUNG AND OLDER ADULT MEN

		Confidence Level				
		Very Sure	Quite Sure	Moderately Sure	Slightly Sure	Not Sure at All
Young men	Mean	7.92	20.36	28.82	28.26	24.62
(N = 225)	*SD*	9.40	7.36	6.72	10.92	17.11
Older men	Mean	14.46	18.69	27.60	29.05	27.02
(N = 65)	*SD*	10.08	8.15	7.77	8.89	11.22
	t	4.67	n.s.	n.s.	n.s.	n.s.
Two-tail p <		.001	—	—	—	—

TABLE 2
JUDGMENT EXTREMITY AT EACH CONFIDENCE LEVEL FOR YOUNG AND OLDER ADULT WOMEN

		Confidence level				
		Very sure	Quite sure	Moderately sure	Slightly sure	Not sure at all
Young women	Mean	6.06	20.74	31.02	32.87	29.86
(N = 132)	*SD*	8.47	9.76	6.67	7.25	14.53
Older women	Mean	11.01	17.99	28.81	30.06	29.44
(N = 89)	*SD*	11.40	7.57	6.69	8.22	11.42
	t	3.79	2.35	2.40	2.60	n.s.
Two-tail p <		.001	.02	.02	.01	—

posts when *S last* said he was "sure" that he could get the car to clear the posts. In the case of each of these three measures, the larger the number of inches, the greater *S*'s subjective probability of failure.

In addition, a performance measure was obtained: namely, the number of inches separating the posts in the second performance series after which the car struck a post three times in a row. The first performance series was considered a practice run in order to give all *S*s some familiarity with the motor aspects of the task.

The subjective probability of failure procedure was administered to the older male and female samples. It was in the case of these two samples, therefore, that relationships between disutility of failure and subjective probability of failure could be examined.

RESULTS AND DISCUSSION

Extremity Index

Tables 1 and 2 present the mean extremity of estimate for items at each of the five confidence levels. For both men and women, highly significant and similar age differences obtain when judgments are "very sure," in the direction of greater extremity for young *S*s than for old.[5] One of the important consequences of aging appears to be a greater unwillingness to "go out on a limb," even though very certain of one's judgment.

A further examination of Tables 1 and 2 indicates a sex difference in the effects of age under confidence conditions lower than

[5] All t tests in this study were two-tailed and carried out by the method described in Walker and Lev (1953, p. 157), which takes account of any heterogeneity of variances.

"very sure." No significant differences in extremity are found between young and older adult males for items placed in other than the "very sure" confidence category. For females, on the other hand, a reversal is obtained under moderate confidence conditions ("quite sure," "moderately sure," and "slightly sure")—the older women now appearing more extreme than young women in their judgments. The magnitudes of the differences in this reversal are not large, but the consistency of these statistically significant effects across the three middle confidence categories renders a sampling error interpretation most unlikely.[6]

In a previous paper (Wallach & Kogan, 1959), it was found that young women were *more* extreme than young men under high confidence conditions in the present judgment situation, while young women were *less* extreme than young men under moderate and low confidence conditions. We attempted to account for this phenomenon in terms of a "counterphobic" release conceptualization, suggesting that female tendencies to take an extreme stand when highly confident might be a reaction against the more restrictive behavioral sanctions generally applied by parents to female than to male children. When a woman achieves old age, we might expect such "counterphobic" tendencies to be partially attentuated. This is what seems to be reflected in the flattened extremity function across the various confidence categories for older females (Table 2). It is as if, under moderately ambiguous circumstances, the older woman now can afford to be a bit more extreme than in her earlier years.

It should be noted, however, that some residue of the sex difference in extremity observed at college age still is manifest in the older sample: women continue to be more extreme than men under high confidence conditions ($t = 1.98$, $p < .05$). None of the extremity differences between older men and older women at moderate and low levels of confidence are significant.

In sum, judgment extremity under high confidence decreases from young adulthood to old age for each sex. Judgment extremity under moderate confidence increases from young adulthood to old age for females. Older women are more extreme than older men under high confidence, as also was the case in comparing younger men and women. Finally, while younger men were more extreme than younger women under moderate and low confidence, no such sex differences were obtained for the older *Ss*.

Confidence Index

Recalling that a smaller index indicates higher over-all confidence, we find the means to be 2.83 ($S.D. = 0.57$) for young men, 3.19 ($S.D. = 0.71$) for older men, 3.11 ($S.D. = 0.52$) for young women, and 3.08 ($S.D. = 0.66$) for older women. With sample sizes of 225, 65, 132, and 89, respectively, these results indicate that young men are significantly higher in over-all confidence than young women ($t = 4.73$, $p < .01$), and young men also are significantly higher in confidence than older men ($t = 3.75$, $p < .01$). On the other hand, there is no significant confidence difference between older men and older women ($t = 0.98$), nor between young women and older women ($t = 0.36$). In sum, there is a sex difference at the young adult level in the direction of greater confidence for men, and also an age difference for men in the direction of less confidence for older males. No such age difference was found for females. These results suggest, then, that the aging experience for males in our culture may be particularly conducive to a decline in confidence level.

Deterrence of Failure Index

An odd-even reliability coefficient was determined for the deterrence of failure index using the Spearman-Brown formula. This reliability coefficient, computed for each sample separately, was .53 for young males,

[6] To evaluate the possibility of sex or age differences in extremity of estimate for particular items, an item analysis was performed comparing each item's mean extremity of estimate for young and old men, young and old women, young men and women, and old men and women. None of the item differences was significant, thus indicating that the obtained sex and age differences in extremity of estimate are not due to specific item content.

TABLE 3

OVER-ALL SCORE AND ITEM COMPARISONS, DETERRENCE OF FAILURE INDEX FOR YOUNG AND OLDER ADULT MEN

		Over-all Score	Item Comparisons											
			1	2	3	4	5	6	7	8	9	10	11	12
Young men	Mean	5.82	5.09	7.00	5.97	5.43	7.72	5.11	4.96	6.67	4.25	5.89	4.33	7.42
(N = 225)	*SD*	1.75	2.30	2.45	2.63	2.61	2.44	2.42	2.83	2.70	2.84	2.49	2.55	2.35
Older men	Mean	6.38	6.03	6.82	8.62	4.82	8.38	4.71	5.49	6.92	5.28	6.97	5.06	7.45
(N = 65)	*SD*	1.31	2.43	2.55	1.63	2.85	1.93	2.78	3.07	2.89	3.14	2.51	2.50	2.25
	t	2.80	2.85	n.s.	10.19	n.s.	2.20	n.s.	n.s.	n.s.	2.34	3.00	2.03	n.s.
Two-tail $p <$		.01	.01	—	.001	—	.04	—	—	—	.03	.01	.05	—

TABLE 4

OVER-ALL SCORE AND ITEM COMPARISONS, DETERRENCE OF FAILURE INDEX FOR YOUNG AND OLDER ADULT WOMEN

		Over-all Score	Item Comparisons											
			1	2	3	4	5	6	7	8	9	10	11	12
Young women	Mean	5.88	5.27	6.93	7.14	5.94	7.32	4.81	4.94	5.89	5.31	6.05	4.31	6.63
(N = 132)	*SD*	1.15	2.14	2.22	2.74	2.27	2.37	2.32	2.61	2.97	2.73	2.66	2.07	2.56
Older women	Mean	6.36	6.89	6.18	8.62	5.35	8.27	4.78	6.24	6.60	5.08	6.42	4.43	7.48
(N = 89)	*SD*	1.61	2.34	3.22	1.83	3.00	2.21	2.89	2.82	3.13	3.24	2.91	3.03	2.20
	t	2.43	5.40	n.s.	4.63	n.s.	3.17	n.s.	3.51	n.s.	n.s.	n.s.	n.s.	2.66
Two-tail $p <$		.02	.001	—	.001	—	.01	—	.001	—	—	—	—	.01

.63 for young females, .80 for older males, and .80 for older females. The internal consistency of the instrument therefore is satisfactorily high. That this consistency increases with age suggests that the degree of disutility of failure becomes less situation-specific and more generalized as a person grows older.

The first columns of Tables 3 and 4 compare young and older *Ss*, within sex, in the over-all deterrence of failure score. A significant difference is obtained between young and older males ($p < .01$), and between young and older females ($p < .02$). In each case, the older *Ss* show a higher disutility or deterrence of failure than the younger. It should further be noted that there were no significant item differences in the direction of lower disutility of failure for older *Ss*. In addition, within the group of older women, there is a correlation of .33 ($p < .01$) between age and the over-all deterrence of failure score. Holding vocabulary score constant, the partial correlation between age and deterrence of failure within the group of older women is .31 ($p < .01$). No relationship obtains, however, within the group of older men ($r = .05$).

Since the older male and female groups have similar age means and variances, these results suggest that the relationship between disutility of failure and age is maintained at higher age levels for women than men. Men, in other words, seem to reach a plateau such that further increases in age are no longer associated with greater disutility of failure. This plateau effect did not occur for the women. These findings suggest that increases with age in the disutility of failure may be more gradual and continuous for women over the lifespan, as opposed to a more abrupt process for men. Possibly, this abruptness is

associated with retirement or the realization of its imminence, factors that are likely to be of little relevance in the lives of most older women.

A comparative analysis of responses to specific items on the deterrence of failure questionnaire yields significant age differences in several content areas (see Tables 3 and 4). Items 1, 3, and 5 show age differences common to both sexes in the direction of higher disutility of failure for older *Ss*. Inspection of the content of these three items reveals that they all are concerned with a choice between modest financial gains as a "sure thing" and very substantial financial gains with the risk of loss. That older *Ss* demand a higher probability of success on these items before choosing the more risky alternative (i.e., have a higher disutility of failure for that alternative) may be a reflection of current financial anxieties. Older individuals probably are more concerned with stabilizing their income in an expanding economy than with the pursuit of a "windfall." Nevertheless, the results obtained come as a bit of a surprise in view of oft-repeated statements by contemporary commentators and essayists (e.g., Whyte, 1956) to the effect that present-day American youth are security-minded in contrast to the more individualistic entrepreneurial bent of a generation or two ago. Given the projective nature of the present items ("Imagine you are advising . . ."), one might have expected such generational value differences to be reflected in the choices made. The present set of findings, then, poses important questions regarding the relative dominance of situational and ideological determinants in decision making. It would appear that the current reality world—financial insecurity in the aged, for example—is capable of bringing about a cognitive restructuring of the value system. Analogous observations in the political sphere have recently been reported by Cantril (1958).

Consider next those items that are sex specific. For males, items 9, 10, and 11 are age differentiating; for females, the relevant items are 7 and 12. In all cases, the older *Ss* show a higher disutility of failure than the younger.

The male items involve risk of death (9) and professional failure (10 and 11). It is difficult to explain why death risk differentiated only for males, unless one resorts to psychoanalytic interpretation emphasizing "potency" symbolism, or to the greater reality of the prisoner-of-war situation for males than females. On the other hand, risk of professional failure clearly is of greater relevance for men than women. It seems reasonable to assume that the older males of the present study have experienced some dissatisfaction in their vocational life, and, hence, show a higher deterrence of failure with regard to advocating an adventurous career choice to others. While it is entirely possible that vocational dissatisfaction is just as likely to be the outcome of an excessive concern for security as of a predisposition toward risk-taking, the older males advise that one settle for less. Here again, we find no evidence in old *Ss* for greater adherence to values embodying risk elements. More recent age-related experiences, possibly including retirement or its expectation, would seem to exert a dominant influence.

The items differentiating age groups exclusively for women concern the risk of defeat in a game situation, and the risk of an unsuccessful marriage. In both cases, the older women show a higher deterrence of failure than the younger. The age difference observed for the "game" item is difficult to interpret. In the case of the "marriage" item, on the other hand, it is not at all surprising that the younger female *Ss* would manifest a lower disutility of failure in the pursuit of so major a goal. The older women, viewing the problem from their distinctive perspective, can well afford to show a higher deterrence of failure in this area.

In sum, the disutility of failure increases from young adulthood to old age for each sex. The function may well be monotonic for females throughout the adult age range, whereas, for males, age differences within the upper range bear no consistent relation to the utility index. In relation to the disutility of failure, then, growing old may be a more continuous process for women and a more abrupt one for men.

TABLE 5

CORRELATIONS OF THE DETERRENCE OF FAILURE INDEX WITH THE SUBJECTIVE PROBABILITY OF FAILURE INDICES (AND PERFORMANCE CONTROL) FOR THE OLDER SAMPLES

		Over-all Deterrence of Failure			
		Older men (N = 65)		Older women (N = 89)	
		r	Partial r[a]	r	Partial r[a]
Estimation measures, car and posts	Inches where "5/5" last given	.42**	.38**	.19	.11
	Inches where "0/5" first given	.23	.20	.34**	.34**
	Inches where "sure" last given	.37**	.37**	.32**	.29**
Performance measure, car and posts	Inches where error criterion reached	.08	.06	.16	.18

[a] Partial correlations hold vocabulary score constant.
** Significant beyond the .01 level, two-tail test.

Subjective Probability of Failure Index

It will be recalled that the subjective probability of failure procedure was administered to the older samples, and that three estimation measures were derived from this task. For the number of inches where "5/5" was last given, "0/5" first given, and "sure" last given, the male means (and associated *S.D.*'s) are 8.32 (3.06), 2.63 (1.53), and 6.25 (2.45), respectively. The corresponding values for the females are 9.61 (3.54), 3.64 (2.29), and 7.80 (3.47), respectively. Turning to the performance measure—namely, the number of inches where the error criterion was reached (excluding the criterion trials)—the male mean is 4.28 (*S.D.* = 0.96), and the female mean is 4.49 (*S.D.* = 1.07). The performance measure was correlated against each of the three estimation measures separately for the two sexes. Only one of the six correlations was significant, that between performance and the "0/5" measure for the older women (r = .26, $p < .05$). There is very little relation, then, between estimation of performance and actual performance in this task.

Consider next the question of sex differences on these measures. The difference in means between the sexes has a t of 2.43 ($p < .02$) for the "5/5" measure, a t of 3.16 ($p < .01$) for the "0/5" measure, and a t of 3.23 ($p < .01$) for the "sure" measure, all in the direction of a higher subjective probability of failure for women than men. No significant sex difference was found, however, for the performance measure. While women manifest a higher subjective probability of failure than men on each of the three *estimation* measures, then, the *performances* of the two sexes do not differ.

Relations Among Aspects of Judgment and Decision Making

Extremity and confidence. The data of Tables 1 and 2 provide four independent confirmations of the hypothesized relationship between extremity and confidence of judgments. In the case of all four samples, judgments of which *Ss* are "very sure" are more extreme than judgments of which *Ss* are "quite sure." Judgments at the "quite sure" level, in turn, are more extreme than judgments made with lower degrees of confidence.

Utility and subjective probability. Table 5 presents the correlation within sex for the older *Ss* between each of the subjective probability of failure indices and the over-all deterrence or disutility of failure index. Two of the three subjective probability of failure measures in the car-and-posts task turn out to be correlated beyond the .01 level

with the over-all deterrence of failure score for each sex, while the performance measure in the car-and-posts task is not related to the over-all deterrence of failure score for either sex.

Since vocabulary score was significantly correlated ($p < .05$) with over-all deterrence of failure score for older men and women (higher vocabulary score related to lower deterrence of failure), partial correlations holding vocabulary scores constant were computed between each car and posts measure and the over-all deterrence of failure score. As Table 5 indicates, none of the relationships changed appreciably with vocabulary scores held constant. For both men and women, there is a partial correlation significant beyond the .01 level between the over-all deterrence of failure score and the estimation measure of the number of inches at which *S* last said he was "sure" he could get the car to clear the posts. In addition, one of the two remaining estimation measures correlates beyond the .01 level with over-all deterrence of failure for each sex, but a different measure for men than for women. For men, the number of inches at which *S* last predicted success in five out of five attempts has a partial r of .38 with over-all deterrence of failure; while the number of inches at which *S* first predicted success in zero out of five attempts shows a nonsignificant trend in the same direction, the partial r with over-all deterrence of failure being .20. For women, the first of these partial r results was a mere .11; while the second partial r was .34, which is highly significant. In addition, another subjective probability measure—namely, the number of inches where "sure" is last given—correlates highly significantly with over-all deterrence of failure for both sexes. Further, the actual performance of each sex on the car and posts procedure shows no relationship to the over-all deterrence of failure score.

The slight difference between men and women in the above results becomes understandable if one considers the sequence of the three estimation measures. The "sure" measure is the last of the three, and hence probably represents a stage of the procedure where most *S*s fully understand the task. If, as is reasonable to expect, men excel women in speed of coming to understand a mechanical task and/or their initial degree of confidence in this task, then the "sure" measure may be the most appropriate of the three estimation scores for evaluating the subjective probability of failure for both sexes. This measure in fact is significantly correlated with the over-all deterrence of failure score for women and for men. Further, we would expect the "5/5" measure, which comes first in the sequence, to be least appropriate for evaluating the women. This measure, in fact, is not related to over-all deterrence of failure for the women, in contrast to the significant relationship obtained with the subsequent "0/5" measure. For men, on the other hand, the "5/5" measure is significantly related to over-all deterrence of failure, and the "0/5" measure shows a trend in this direction.

If the "5/5" measure is not reflecting subjective probability of failure for the women because this measure occurs before women fully understand or have gained confidence regarding the task, then sticking to a "5/5" response might well be a simple persevering tendency. That is, women might tend to keep saying "5/5" just because it is the response they started out making and they are uncertain what to do next. If this interpretation is correct, we might expect this "5/5" measure to correlate inversely with the confidence index for women. Waiting longer before giving up the "5/5" response (i.e., giving it up only after the posts are closer together) thus would reflect *S*'s general uncertainty over what to do in this task.

The correlation between the number of inches where *S* last responded "5/5" and the confidence index is $-.29$ ($p < .01$) for women, indicating that perseveration in the "5/5" measure is related to lower confidence. This correlation is not significant for the men. In addition, correlations between each of the other two estimation measures (both of which occur later in the task) and the confidence index are not significant for women or men.

Since a significant correlation was ob-

tained for women between the "0/5" estimation measure and performance, it was necessary to control the latter in evaluating the relationship between that estimation measure and the deterrence of failure score. The partial correlation coefficient holding performance constant was .32 ($p < .01$), indicating that the relationship between the "0/5" measure and deterrence of failure is not an artifact of ability differences.

We may conclude, therefore, that disutility or deterrence of failure is positively related to subjective probability of failure for the two samples where data are available on this question. The presence of such a relationship between utility and subjective probability in decision making provides a warrant for thinking in terms of greater or lesser caution as a dimension of personality. "Caution" as used here thus refers to the presence of both high deterrence (or disutility) of failure and high subjective probability of failure.

Relations across domains. Finally, we turn to the question of possible relations between utility and subjective probability, on the one hand, and confidence and extremity of judgment, on the other.

We first consider extremity. Three indices of subjective probability of failure were available for the older sample of each sex, as also was a measure of extremity for decisions made at each of five confidence levels. Of the 30 resulting correlations between a subjective probability measure and an extremity measure obtained for each sex separately, only five were significant, three in the predicted direction and two in the opposite direction. In addition, the over-all deterrence of failure score could be correlated with each of the five extremity scores for each of the four age and sex samples. Of the 20 resulting correlations, only one was significant, and in the unpredicted direction. Although one might expect greater judgment extremity to be related to lower disutility of failure and lower subjective probability of failure, the present research hence yields no evidence in support of that expectation.

Turning to confidence, of the six correlations obtained between a subjective probability measure and the over-all confidence index, only one was significant. Further, the over-all deterrence of failure score was unrelated to confidence of judgment in all four samples. Although one might expect higher judgment confidence to be related to lower subjective probability of failure and lower disutility of failure, the present results again do not support that expectation.

It is evident, then, that the trend toward conceptual convergence of the domains of judgment and decision making is without empirical basis for the particular measures employed in the present study. These findings do not, of course, rule out possible relationships between other judgmental and decision indices. Until further careful empirical work is carried out in these areas, however, attempts to conceptualize various judgmental and cognitive processes in decision and risk terms must be considered premature.

REFERENCES

Basowitz, H., & Korchin, S. J. Age differences in the perception of closure. *J. abnorm. soc. Psychol.*, 1957, 54, 93–97.

Birren, J. E. Age changes in speed of simple responses and perception and their significance for complex behavior. In *Old age and the modern world*. London: Livingstone, 1955. Pp. 235–247.

Botwinick, J. Drives, expectancies, and emotions. In J. E. Birren (Ed.), *Handbook of aging and the individual*. Chicago: Univ. of Chicago Press, 1959.

Brim, O. G., Jr. Attitude content-intensity and probability expectations. *Amer. sociol. Rev.*, 1955, 20, 68–76.

Brim, O. G., Jr., & Hoff, D. B. Individual and situational differences in desire for certainty. *J. abnorm. soc. Psychol.*, 1957, 54, 225–229.

Bruner, J. S., & Tajfel, H. Cognitive risk and environmental change. *J. abnorm. soc. Psychol.*, 1960, in press.

Cantril, H. *The politics of despair*. New York: Basic Books, 1958.

Cohen, J. The measurement of uncertainty. *Discovery*, 1958, 19, 286–290.

Cohen, J., Dearnaley, E. J., & Hansel, C. E. M. Risk and hazard: influence of training on the performance of bus drivers. *Operat. res. Quar.*, 1956, 7, 67–82.

Dana, R. H. A comparison of four verbal subtests on the Wechsler-Bellevue, Form I, and the WAIS. *J. clin. Psychol.*, 1957, 13, 70–71.

Doppelt, J. E., & Wallace, W. L. Standardization of the Wechsler Adult Intelligence Scale for older persons. *J. abnorm. soc. Psychol.*, 1955, 51, 312–330.

Feather, N. T. Subjective probability and decision under uncertainty. *Psychol. Rev.*, 1959, 66, 150–164.

Gardner, R. W. Cognitive styles in categorizing behavior. *J. Pers.*, 1953, 22, 214–233.

Hamilton, V. Perceptual and personality dynamics in reactions to ambiguity. *Brit. J. Psychol.*, 1957, 48, 200–215.

Kogan, N., & Wallach, M. A. Certainty of judgment and the evaluation of risk. *Psychol. Rep.*, 1960, 6, 207–213.

Korchin, S. J., & Basowitz, H. The judgment of ambiguous stimuli as an index of cognitive functioning in aging. *J. Pers.*, 1956, 25, 81–95.

Korchin, S. J., & Basowitz, H. Age differences in verbal learning. *J. abnorm. soc. Psychol.*, 1957, 54, 64–69.

Luce, R. D., & Raiffa, H. *Games and decisions.* New York: Wiley, 1957.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. *The measurement of meaning.* Urbana: Univer. Illinois Press, 1957.

Pettigrew, T. F. The measurement and correlates of category width as a cognitive variable. *J. Pers.*, 1958, 26, 532–544.

Robinson, W. S. The statistical measurement of agreement. *Amer. sociol. Rev.*, 1957, 22, 17–25.

Robinson, W. S. The geometric interpretation of agreement. *Amer. sociol. Rev.*, 1959, 24, 338–345.

Ruch, F. L. The differentiative effects of age upon human learning. *J. gen. Psychol.*, 1934, 11, 261–268.

Strother, C. R., Schaie, K. W., & Horst, P. The relationship between advanced age and mental abilities. *J. abnorm. soc. Psychol.*, 1957, 55, 166–170.

Walker, Helen M., & Lev, J. *Statistical inference.* New York: Holt, 1953.

Wallach, M. A., & Caron, A. J. Attribute criteriality and sex-linked conservatism as determinants of psychological similarity. *J. abnorm. soc. Psychol.*, 1959, 59, 43–50.

Wallach, M. A., & Kogan, N. Sex differences and judgment processes. *J. Pers.*, 1959, 27, 555–564.

Wechsler, D. *The measurement of adult intelligence.* Baltimore: Williams & Wilkins, 1944.

Whyte, W. H., Jr. *The organization man.* New York: Simon & Schuster, 1956.

(Manuscript received January 25, 1960.)

A man can understand astronomy only by being an astronomer; he can understand entomology only by being an entomologist (or, perhaps, an insect); but he can understand a great deal of anthropology merely by being a man. He is himself the animal which he studies. Hence arises the fact which strikes the eye everywhere in the records of ethnology and folklore—the fact that the same frigid and detached spirit which leads to success in the study of astronomy or botany leads to disaster in the study of mythology or human origins. It is necessary to cease to be a man in order to do justice to a microbe; it is not necessary to cease to be a man in order to do justice to men. That same suppression of sympathies, that same waving away of intuitions or guesswork, which make a man preternaturally clever in dealing with the stomach of a spider, will make him preternaturally stupid in dealing with the heart of man.

G. K. CHESTERTON, *Heretics.*

In this article, correlation techniques are used to construct a quantitative hypothesis indicating the relative importance of blood lipids, blood pressure, family history of atherosclerotic disease, body build, and psychological traits in the production of coronary heart disease. Responses to the psychological questionnaire characterize patients with coronary disease as accepting of multiple responsibilities, free from neurotic anxiety, self-critical, and possessing an exuberant driving temperament.

THE FACTOR STRUCTURE OF VARIABLES CONCERNED WITH CORONARY ARTERY DISEASE

by Lee D. Cady, Jr., Menard M. Gertler, Lida G. Gottsch, and Max A. Woodbury[1]

New York University

INTRODUCTION

The causes of coronary heart disease are believed to be associated with such factors as blood lipids, blood pressure, family history of cardiovascular diseases, body build, psychological traits—and probably other factors as well. This article reports on a study which used vector methods in tandem with regression techniques to construct a descriptive hypothesis indicating the relative importance of the above-named factors in the production of coronary heart disease.

METHOD

Fifteen variables were considered to be reliable measures of the factors under consideration. In addition to the coronary criterion itself, these variables were: four personality factors (A, F, H, L) from the Cattell Sixteen Personality Factor Questionnaire (Cattell & Stice); three family history items about the subject's siblings, father, and mother; weight; height; Sheldon's mesomorphic index (Sheldon et al., 1940); serum lipid phosphorus; serum total cholesterol; diastolic blood pressure; and systolic blood pressure. The variables were recorded from persons with coronary artery disease without known hypertensive disease, as well as from normal people.

The populations studied were under continuous observation for over ten years. The disease population was composed of white men from 30 to 50 years of age, with a mean age of 38, who were either patients at the Massachusetts General Hospital or members of the consulting staff there. The control population was made up of 146 employees of the Lever Brothers Co., primarily from their Boston office but also some from the New York office, of all occupational levels ranging from executive vice-presidents to shipping clerks and maintenance men. A detailed description of the occupational, educational, and ethnic backgrounds of the subjects is available in a previous study (Gertler, White, et al., 1954).

Clinical and laboratory information was evaluated for 61 patients and 146 control subjects. The Cattell Sixteen Personality Factor Questionnaire was given to 35 patients and an approximately equal number of controls. [The Cattell Questionnaire which is group administered with pencil and paper requires less than 40 minutes to give and is machine scoreable (Cattell & Stice).] Quantitative data, such as those given by questionnaire scores, serum total cholesterol, and systolic blood pressure, were used

[1] The authors are indebted to Howard A. Rusk, Director, Institute of Physical Medicine and Rehabilitation, College of Medicine and Harold K. Work, Director, Research Division College of Engineering, New York University; the IBM Corporation, The National Institutes of Health, and The John A. Hartford Foundation Inc. for creating staffing and financing conditions favoring the conduct of this research.

TABLE 1

CORRELATIONS FOR THE STUDY OF CORONARY DISEASE*

Variables															
Coronary criterion	1														
L	2	−.20													
H	3	.14	−.27												
F	4	−.09	.06	.55											
A	5	.17	.02	.34	.43										
Siblings	6	.32	−.02	−.14	−.09	.02									
Father	7	.26	−.00	.17	.19	.08	.22								
Mother	8	.22	.11	−.06	.18	−.07	.26	−.05							
Weight	9	−.08	.03	−.06	.13	−.19	−.09	−.13	.21						
Height	10	−.30	−.07	−.02	−.08	−.14	.00	−.11	−.09	.41					
Mesomorphy	11	.25	.16	−.08	.08	−.08	−.04	−.05	.16	.46	−.11				
Lipid phosphorus	12	.11	.06	.15	.33	.27	.15	.10	.29	.10	−.10	.24			
Cholesterol	13	.50	−.11	.06	.20	.17	.10	.02	.07	.02	−.15	.10	.63		
D. Blood pressure	14	.23	.03	−.12	.08	−.07	−.20	−.19	.16	.17	.05	.18	.07	.25	
S. Blood pressure	15	.24	.06	−.09	.08	.01	−.04	.01	.17	.08	.05	.04	.12	.15	.75
		1	2	3	4	5	6	7	8	9	10	11	12	13	14

* A correlation coefficient of 0.16 is at the 5% level of significance for a sample size of one hundred and fifty.

directly in this form to obtain the correlations presented in Table 1. Dichotomous data, such as the presence of coronary disease in the subject, and atherosclerotic disease in the mother, father, or siblings, were each given scores of one if present and zero if absent.

The relations between all of the 15 variables in all possible combinations taken two at a time are expressed by the 105 correlation coefficients presented in Table 1. A principal component factor analysis was applied to the matrix of correlations to gain a better understanding of the complex relationship between the 15 variables, and to resolve these relations into a simpler three-dimensional factor structure characterizing the 15 variables related to coronary heart disease (See Anderson, 1958). Unities were used as diagonal elements of the correlations matrix and no rotation of axes was made. The correlations were used to derive discriminant function Equations 1 and 2. These equations assign weights to be applied to the psychological scores A, F, H, and L, and other variables, for the purpose of discriminating between coronary and control groups on the basis of psychological and other traits.

RESULTS

Inspection of the correlations in Table 1 indicates that the cholesterol variable is the variable most closely related to coronary disease. The blood lipid variables are so closely related to each other as to be similar measures of the same thing. As expected, height and weight appeared definitely correlated with each other as did systolic and diastolic blood pressures.

Inspection of the simplified factor structure of the 15 variables in Table 2 reveals that factors one, three, and five are most clearly related to coronary heart disease in the normotensive population of men of ages 30 to 50 and that factors two, four, and six are not related to coronary heart disease in this population. Eighty-five per cent of the total variation in these 15 variables was accounted for by six mutually independent factors. Each factor had components oriented with respect to the 15 variables. One

TABLE 2
THE FACTOR STRUCTURE OF FIFTEEN VARIABLES

Factor		1*	2	3*	4	5*	6
% of Total Variance		32%	21%	14%	10%	6%	4%
Eigen Value		27149	22093	17865	14815	12148	11326
Variables							
Coronary criterion	1	.37*	−.02	.41*	.07	.21	.03
L.	2	.00	−.12	.06	−.07	−.78*	−.03
H.	3	.16	.40*	−.40*	.07	.18	−.11
F.	4	.30*	.25	−.43*	−.05	−.25*	−.14
A.	5	.23	.36	−.14	.15	−.15	.02
Siblings	6	.12	.07	.43*	−.33	.09	−.43*
Father	7	.12	.27	.19	−.07	.03	−.39
Mother	8	.26	−.18	.08	−.30	−.24	−.23
Weight	9	.08	−.36*	−.35	−.42*	.11	−.06
Height	10	−.16	−.20	−.31	−.14	.25*	−.49*
Mesomorphy	11	.22	−.25	−.09	−.35	.09	.42*
Lipid phosphorus	12	.43	.08	−.08	−.23	−.06	.10
Cholesterol	13	.44*	.25	.09	.03	.24*	.22
D. Blood pressure	14	.27*	−.42*	−.11	.45*	.01	−.07
S. Blood pressure	15	.27	−.33*	−.03	.45*	−.05	−.32

* An important factor or factor component.

and three are the two most closely related to coronary disease. The three "coronary" factors account for 44 per cent of the variation in these variables and the factors have large lipid, psychological, and family history components. Factor five was a psychological factor that was weakly and inversely related to coronary disease. These data showed no tendency for the coronary group to be overweight when mesomorphic build was taken into account. The magnitude of the weight component of only factors two and four was high and neither of these factors had a close relationship with coronary disease. Factors two and four were oppositely directed with respect to blood pressure, while factor two had a large psychological component and factor four had a large family history component. The sixth factor was almost exclusively a body build factor with major components of height and mesomorphy and no relation to coronary disease.

The terms of the discriminant function Equation 1 were arranged in order of the importance of the variables in coronary artery disease, when the variables were measured in standard score form.

$$\begin{aligned}\text{Coronary Status} = {} & 0.80X_1 - 0.65X_2 \\ & - 0.58X_3 + 0.43X_4 \\ & + 0.41X_5 + 0.39X_6 \\ & + 0.33X_7 + 0.33X_8 \\ & + 0.23X_9 - 0.08X_{10} \\ & - 0.07X_{11} - 0.06X_{12} \\ & - 0.06X_{13}\end{aligned} \qquad (1)$$

where variables X_1 through X_{13} are: (1) serum total cholesterol, (2) serum lipid phosphorus, (3) psychological Factor F, (4) maternal history of cerebral or coronary artery disease, (5) mesomorphic index, (6) psychological Factor H, (7) paternal history of cerebral or coronary artery disease, (8) psychological Factor A, (9) systolic blood pressure, (10) weight, (11) height, (12) diastolic blood pressure, and (13) psychological Factor L. Serum lipids were clearly more important than the other variables in this group of normotensive men; blood pressure was less important than family history, psychological factors, and body build.

For raw scores, the discriminant function Equation 2 indicated that psychological

variables are related to coronary artery disease.

$$\text{Coronary Status} = 0.68\ A - 0.62\ F + 0.27\ H - 0.43\ L \quad (2)$$

(Dr. Raymond Cattell kindly rendered interpretation of this equation.) Taken together, the psychological traits which characterized the coronary subjects were: (1) A constitutional cyclothymic temperament (as opposed to schizothymic); and (2) An environmentally induced desurgency manifested by: (a) inhibitory habits with respect to emotional expression; (b) general concern about life, with a self-critical (stressed) acceptance of responsibilities to be shouldered; and (c) freedom from neurotic or floating anxiety and withdrawal.

It has been found in twin studies that psychological Factors A and H have a highly inherited constitutional element. These two factors together comprise the cyclothymic temperament. Factor F reflects the degree of punishingness of the environment, the acquisition of inhibitory patterns of emotional expression and is somewhat opposed to Factor L which represents paranoic tendencies to project one's faults or difficulties onto others. The effort stress reaction associated with fast pulse rate, elevated serum levels of corticoids, creatinine, arginine, valine, methionine, glycine, serine, and low serum glutanic acid is greater in persons who are high in Factors A, H, and F, but low in L (Cattell & Stice).

DISCUSSION

There is an overlap of psychosomatic patterns underlying different diseases. Some of the same personality characteristics can be observed both in patients with coronary artery disease and in hypertensive patients. Furthermore, coronary disease is frequently associated with hypertension in the same individual. Hypertensive patients can be characterized by a difficulty in self-assertiveness and expression of resentment; they control hostility, but are bitter about it underneath (Kelman, Alexander, & Stein, 1958). Further study is needed to relate the self-criticism of the coronary patient to these expressive difficulties experienced by hypertensives.

Psychosomatic interactions are complex and probably but a few of these are pointed to in this study. Clearly, personality factors can be used as part of a battery of measures to distinguish between patients with different diseases. Studies of the processes of psychosomatization in coronary artery disease are hindered by the necessity of studying persons after they obviously have the disease and by an uncertainty about possible undetected arterial disease in apparently healthy control subjects. A discriminant function capable of accurately estimating relative proneness to coronary disease could overcome this hindrance so that separation of extremely prone or nonprone persons becomes possible with such a device.

The limitations in sample size and the biases attending the use of patient records from a large medical center and from prominent consultants impose restrictions on any precise interpretation of the data presented. For these reasons, the results of this analysis are reported as an hypothesis to be tested more extensively in the future.

Such relationships as elevated serum lipids, positive family history of cardiovascular diseases, psychological traits, were considered as the dimensions of the problem of coronary heart disease. These variables or dimensions are interrelated (correlated) with each other and are not independent of each other (i.e., they are not orthogonal or vectorally perpendicular). The nonorthogonal system of co-ordinates was resolved into a simpler system of independent co-ordinates with three orthogonal dimensions related to coronary heart disease.

It was anticipated that the statistically determined factors would be identifiable as some real factors such as blood lipids, body build, etc. The factor loadings are the basis for such identifications. Each factor loading is a measure of the correlation between a variable and a factor, since a correlation coefficient can be interpreted geometrically as the cosine of the angle between the line of action of two variables. It is also a measure of the difference in the directions of force of a variable and a factor. A factor loading of +0.4 is interpreted as large for these data and the force of the component variable is

directed along a line which had an angular cosine of +0.4 with the factor.

Inspection of the factor loadings for the important variables of each factor serves to indicate the possibility of a common causal association for blood lipid elevations and certain psychological traits, as well as for blood pressure elevations and positive family history of cardiovascular diseases.

CONCLUSION

There appears to be a statistical relationship between proneness toward coronary artery disease and certain measurable variables. These results will be of help in the design of clinical trials of preventive measures to delay and forestall the initial coronary episode. Effort is being exerted to test the accuracy of individual prediction based on these retrospective studies. Accurate individual estimation of proneness to coronary heart disease would make preventive therapy highly selective and realistic.

REFERENCES

Anderson, T. W. *Introduction to multivariate statistics.* New York: John Wiley & Sons, 1958.

Cattell, R. B., & Stice, G. F. *The 16 personality factor questionnaire.* Champaign, Illinois: Institute for Personality and Ability Testing, 1602 Coronado Drive.

Gertler, M. M., White, P. D., et al. *Coronary heart disease in young adults.* Cambridge: Harvard University Press, 1954.

Kelman, C., Alexander, F., & Stein, M. I. *Some psychological correlates of six somatic disturbances.* Paper read at American Psychological Association, Washington, D. C., August, 1958.

Sheldon, W. H., et. al. *The varieties of human physique.* New York: Harper Brothers, 1940.

(Manuscript received March 25, 1960.)

Thus, if we take a very broad view of the history of science, we may distinguish in it four main phases. The first is the empirical development of Egyptian and Mesopotamian knowledge. The second is the building of a rational foundation of astounding beauty and strength by the Greeks. The third, and until recently the least known, is the mediaeval period—many centuries of groping. Immense efforts were spent to solve pseudo-problems, chiefly to conciliate the results of Greek philosophy with religious dogmas of various kinds. Such efforts were naturally sterile, as far as their main object was concerned, but they brought into being many incidental results. The main result was the incubation of the experimental spirit. Its final emergence marks the transition between the third period and the fourth, which is the period of modern science. Note that out of these four periods the first is entirely Oriental, the third is mostly but not exclusively so; the second and fourth are exclusively Western.

GEORGE SARTON, *History of Science and the New Humanism*

The selection of a particular null hypothesis to be tested in a clinical situation, where only small numbers of homogeneous patients are available, often leads unwittingly to bias in its favor. It is frequently forgotten that failure to obtain a statistically significant result is equivalent to a nondecisive verdict, not to a verdict to reject a hypothesis (null or otherwise). Here the shortcomings of "extensive" models (in which the degrees of freedom do not exceed the number of patients in the sample) are discussed and a shift to an "intensive" model is indicated, based on the acceptance of a general underlying degree of variability in day-to-day behavior and affect.

STOCHASTIC MODELS OF THE SINGLE CASE AS THE BASIS OF CLINICAL RESEARCH DESIGN

by Jacob B. Chassan

*Saint Elizabeth's Hospital, Washington, D. C.**

INTRODUCTION

THE purpose of this paper is, first, to examine critically certain aspects of current conventional clinical research methodology, in particular those aspects which have to do with statistical inferential processes within the framework of formalized clinical design and, second, to present an alternative approach based on a stochastic view of psychopathology and interpersonal processes in general. The conventional approach will be identified with the type of study in which the degrees of freedom for the application of statistical tests do not exceed the number of patients under study. It will be referred to as the *extensive* model. The statistical techniques which are used in the extensive model are limited essentially to end-point observations. Among other limitations, it will be seen that the model offers little or no chance of a steady progress from one study to another, or for planning a well-integrated program of clinical research on the ward in which the results of one study indicate the possible directions of the next, etc.

The approach to clinical design which will here be argued as overcoming the practical limitations and the essential weakness of extensive design will be the concept of design based on the *intensive* statistical study of each patient over time. The paper will deal essentially with the problems of clinical design in the evaluation of treatment on the ward in the setting of the mental hospital, although the validity of the intensive approach in other areas, such as the study of individual psychotherapy, will be apparent.

THE EXTENSIVE MODEL

Limitation on size of the population as a source of implicit bias and insufficiency

Perhaps the greatest difficulty of the extensive model stems from the relatively small number of sufficiently homogeneous patients in a mental hospital at any given time who can be brought together for the purpose of a specified piece of clinical research. Even in a large hospital with thousands of patients, variability in age, sex, length of hospitalization, socioeconomic level, physiological characteristics, etc., limit to a considerable extent the selection of patients for any given study. Other factors, such as medical counter-indications and management problems, also play a part in curtailing the size of the population (N).

The principal standard statistical inferential procedure for comparing the difference in results obtained between two

* Now with the Office of Education, Department of Health, Education and Welfare, Washington 25, D. C.

groups is first to state a *null hypothesis*, and to apply an appropriate statistical test to the data. If the test reveals that the results are incompatible with the truth of the null hypothesis at some stated level of significance, it is rejected. Otherwise, the results are considered as consistent with the null hypothesis, and the latter, at least in effect, is accepted as true.

Now it is generally true that the same set of results can be consistent in this sense with any number of hypotheses selected a priori *as* the null hypothesis and such that these hypotheses are not consistent with one another. However, as long as one is dealing with large samples (and the respective parameters defining the competing hypotheses are sufficiently disparate), it may not matter in practice which of two or more competing hypotheses is selected as the null hypothesis. But when N is relatively small, the selection of a particular null hypothesis often represents an unwitting bias in its favor. This appears to be a common occurrence in the application of statistics to clinical research. It is not a consequence of the statistical methodology itself, but rather the result of misinterpreting the implications of a result which is not statistically significant. What often seems to be forgotten is that such a result is essentially an inconclusive result.

The acceptance of the null hypothesis as true for failure to obtain a statistically significant result leads to a bias in favor of the acceptance of the particular null hypothesis selected for testing. Most such biases are essentially conventional rather than personal ones because it is an almost universal practice to select the null hypothesis which states that there is no true difference in effect. This is none the less a serious bias, for it consistently operates *against* an active treatment, when the latter is tested against a control. While it is logically just as valid to test a hypothesis which assumes a true underlying difference, it should be noted again that the failure to obtain statistical significance for the rejection of such a hypothesis does not establish it as true either. The testing of such null hypotheses, as opposed to the more conventional one, could easily lead to the replacement of conventional bias by personal bias, and could therefore hardly be an improvement.

A somewhat painstaking analysis of the problem of sample size in relation to the detection of drug effects by the use of the extensive model is given by Kramer and Greenhouse (1959). The tables which they present on the numbers of cases required to detect true percentage differences in effect, together with their presentation of relevant descriptive statistics concerning the size and certain of the broad characteristics of schizophrenic patient populations in various of the state hospital systems, provide further convincing evidence of the insufficiency of the extensive model. As a means of increasing N, they suggest the pooling of data from various hospitals. However, the attempt to pool data within the structure of formal designs would encounter many difficulties. Besides specific differences in research interests from one institution to another, there are differences in point of view concerning the definition of various aspects of psychopathology, as well as variability in administrative and clinical practices, all of which can place heavy burdens on the structure of formal interhospital clinical designs. Further, one must consider that tensions may increase to the detriment of the research as a whole, as the distance increases from the actual operation of the design on the ward to the geographical point from which the research is directed. Even if it were otherwise feasible to develop interhospital clinical designs, the total volume of clinical research organized in this way, in terms of the number of different designs that can extend over essentially the same period to satisfy the varied interests of clinical investigators, would be extremely limited as a result of the large expenditure on the available N for even a single design, or at best, for just a small number of distinct, interhospital designs.

Methodological biases resulting from contagion

A form of bias against the detection of true differences in results stems from a

generally unwarranted assumption of statistical independence of the observations concerning the behavior and affect of patients who are on the same ward. The existence of positive correlation between patients, together with the failure to take this into account in common tests (as are used in the extensive model) for significant differences between groups, results in understating the true level of significance corresponding to an observed difference in the means. For example, in the case of a normally distributed variable, if z_o represents a critical ratio on the assumption of independent observations, then $z_r = z_o/\sqrt{1-r}$ is the critical ratio corresponding to the same observed difference in means when r is the appropriate coefficient of correlation which describes or defines the degree of statistical dependence between paired patient observations on a given variable. As long as r is positive, $z_r > z_o$, and the use of z_o instead of z_r involves an understatement of the critical ratio by a factor of $(\sqrt{1-r})^{-1}$. In practice, r cannot be estimated from the data of the extensive model. The use of standard critical ratios implicitly assumes that $r = o$, contrary to the reasonable expectation of at least some contagious or communicable components of the variability of behavioral and affective states. Any attempt to eliminate the difficulty by separating the active treatment from the control group by using two different wards will obviously result in a confounding of treatment and ward effect. In such a design, the presence of positive correlation within groups on the basis of a contagious effect and zero correlation across groups (as such a separation would presumably entail) would result in the opposite bias of overstating levels of significance.[1]

Questions of operational significance

An essential point about the extensive model which has been stressed thus far is the difficulty of obtaining a sufficiently large number of patients to establish an observed difference in effect as statistically significant, when a true difference does "in fact" exist. In this section the argument against the extensive model will be carried further to cover the case in which a statistically significant difference does occur.

The essential difficulties here concern the theoretical population of patients from which the patients selected for a study are considered as a random or a representative sample. While it may not be too difficult to *describe* the theoretical population or statistical distribution simply as an extension of the design sample, the operational or practical implications of a statistically significant result within the framework of extensive design are vague, or at best limited. Even in a design with a relatively small number of patients, the number needed for a statistically significant result is such that a considerable extent of variability in patient parameters and background characteristics will be found among the patients. For example, if one wishes to develop a clinical design even within a relatively stable part of a hospital's patient population, one usually has to be satisfied to include patients whose age and period of hospitalization may vary from 5 years to 15 years, whose symptoms may differ to a considerable extent though they may all be classified within the same broad diagnosis, and whose various physiological characteristics may also vary extensively. Neither can variation in factors descriptive of early development, family structure and childhood experience, marital status, occupation, intelligence, educational level, treatment history, etc., be generally kept within reasonably narrow limits. Such variation can easily result in a level of significance which, again, in consideration of the necessary limitations on the number of patients in a study, easily can be more a product of an imbalance of relevant characteristics between the treatment and control group than a true measure of the statistical significance of the effect which the design has been formulated to study.[2]

[1] The relationship between the correlated and uncorrelated critical ratios would then be given by $z_r = z_o \Big/ \sqrt{1 + r\left(\frac{N}{2} - 1\right)}$.

[2] The process of pairing patients in the extensive model is one which in practice can achieve a balance with respect to only a very small number of possibly relevant characteristics, and then the

Further, even if one can assume that one has obtained a statistically significant difference within an extensive design, free of either apparent or hidden bias, the variations in patient background characteristics impose severe limitations on the specificity of application of the significant finding. The difficulty here follows from the realization that a statistically significant difference between two groups of patients within the framework of an extensive model, in general, provides little, if indeed any, data at all concerning which of the specific patient characteristics and parameters, and/or their combinations, may have a bearing on the over-all demonstrated effect. For example, one cannot distinguish within an active treatment group which of the patients "improved" as a result of treatment on the drug from those who improved (within the same active group) in the same way that patients in a matched placebo or controlled group may have improved. Thus one cannot, in general, tie the treatment effect in any way to patient background characteristics and parameters. From an operational point of view, insofar as the results of a study are to be considered as data for subsequent clinical decisions concerning the treatment of patients, it follows that even though a statistically significant result demonstrating an effect is obtained, one still has a minimum of data for refinement in the selection of patients for treatment. Similarly, it seems highly improbable (under the limitations of the extensive model discussed thus far) that an organized program of clinical research can be developed on the basis of a series of extensive designs, such that the results of one study of the series leads logically and with a sense of direction to the formulation of the next, or of reasonably delineated possible nexts. The following comment by Kety (1959) is relevant to the present discussion. He writes:

> "Errors involved in sampling from heterogeneous populations may help to explain the high frequency with which findings of one group fail to be confirmed by those of another. Recognition that any sample of schizophrenics is probably a heterogeneous one would seem to indicate the importance of analyzing data not only for mean values but also for significant deviations of individual values from group values. The biochemical characteristics of phenylketonuria would hardly have been detected in an average value for phenylalanine levels in blood in a large group of mentally retarded patients."

The general lack of operational significance of the extensive model is also apparent from a practical consideration of statistical inferential processes which are relevant to design samples and their theoretical parent distributions. Recognizing that the purpose of selecting and observing a design sample is to make inferences from the sample to its parent population (or distribution), the validity of such inferences must depend upon the randomness of the sample in relation to the parent population. Suppose the subjects chosen for an experimental design are actually selected by some appropriate method of randomization (allowing for such devices as stratifying and pairing) from an existing well-defined group of patients (representative of some "hypothetically infinite population"). Then there can be little question concerning the applicability or extension of inferences drawn from a statistical analysis of the design data beyond the design patients to other patients randomly (or even, let us say, randomly in effect) selected from the same parent population. For example, suppose the data of the design show a statistically significant difference favoring drug A over drug B in the treatment of some condition. Then a decision in favor of drug A is clearly indicated for treatment of the same condition in other patients randomly selected from the parent population. This, of course, follows from the expectation based on the results of the design that on an average a greater percentage of similar patients would improve on A than on B.

Now, if we return to a consideration of what might be called the existential limitations of clinical design, we often find that the same difficulties which are involved in

characteristics selected for pairing are often those for which the data are easiest to acquire and not necessarily those which may be the most relevant. In any event, the idea that all or even most of the important biases can in practice be eliminated by pairing is in all likelihood more a delusion of methodology than a reasonable expectation.

obtaining a well-balanced design with a sufficiently large number of patients are also encountered in the problem of the operational extension of statistically significant results. If within a hospital practically all available patients who meet a set of criteria for a study need to be selected *for* the study to obtain a reasonable sample, then there can be few, if any, patients left who meet the same criteria and who can therefore be considered as representative of the same parent population as the patients of the design study. Consequently, one has few patients left for a statistically valid application of the results of the significant finding of the experimental design, at least within a reasonable period after the conclusion of a study. At best, under the limitations which have been discussed thus far, it is fairly evident that the population of reference corresponding to the extensive model is vague and ill-defined. This conclusion will be of particular importance in relation to the comparison of statistical inferential processes between extensive and stochastic models which will be presented in a later section of this paper.

THE INTENSIVE DESIGN

Stochastic definition of patient-state as the foundation of stochastic design

The shift from the extensive model of design to the stochastic model is based upon the acceptance of a general underlying variability within the subject, or patient, with respect to his day-to-day attributes and characteristics which are thought relevant to the description and understanding of his personality and psychopathology. This variability is the basis for defining the patient-state in terms of statistical distributions, or probabilities estimated from sequences of observations of the same patient,[3] and consequently provides the framework for the application of statistical inference and design to the data of the single patient.

[3] This concept has been discussed by the writer in some detail in previous articles (Chassan, 1957; 1959).

Delineation and control of relevant patient characteristics

Once a significant difference appears within the single case design, one can specify the particular patient background variables and other relevant characteristics in whose presence the significant result was actually obtained. That is, we know the patient's age, his diagnosis, various of his physiological characteristics, and many other of his possibly pertinent characteristics and life's experiences. This is in sharp contrast to the value of a statistically significant result obtained from the extensive model. Furthermore, in a design which is based upon a detailed statistical analysis of each case, each patient serves as his own control, and a statistically significant result cannot be an artifact of a lack of homogeneity between patients, or of poor pairing. In the intensive single case analysis, not only are such factors as age, sex, diagnosis, educational level, and intelligence kept constant, but so are all of the possibly significant interpersonal experiences in life up to the beginning of a study, including childhood experiences, significant figures in early life, sexuality, etc.; all of the relevant physiological and anatomical characteristics (including, e.g., such variables as brain structure and the entire previous history of metabolic functioning); and all possibly relevant genetic factors. This degree of control or anything near it is obviously impossible in the extensive model.

Generalization from the statistics of the single case

An argument which is sometimes stated against the single-case inference is: after all, the result is significant with respect to but a single case. In the extensive model, although statistically significant results are hard to come by because of the dearth of degrees of freedom, once a significant result is obtained in the extensive model, you have something which pertains to a whole population of cases. You have achieved a result more capable of generalization, and therefore a more useful result.

Such an argument favoring the extensive model cannot validly be maintained once its limitations, as discussed in previous sections of this paper, are borne in mind. It is a common fallacy to assume that if a significant effect has been demonstrated by means of an extensive design, the effect exists for every, or for most, or even for an appreciable proportion of the members of the population. Not only is this generally false, but we have seen from our previous discussion that the extensive model does not even permit the identification of patient characteristics accompanying the tested effect. Furthermore, we have seen that the populations themselves corresponding to extensive design samples are vague, particularly as an attempt is made to apply the results toward the subsequent selection of patients for treatment or for directed research effort.

It has been stressed that the vagueness of the population about which inferences from an extensive design sample can be drawn in terms of subsequent applications is one of the weaknesses of the extensive model of design. The opposite argument against the overspecification of the population of the intensive single-case study, to the extent that statistical inferences are strictly valid only with respect to the single case itself, is, of course, technically correct. However, if one considers the intensive design data of the single patient as a function of all of his relevant characteristics (and, if one wishes, a random component), then the statistical population with which the single case is identified can in theory be described or defined in terms of these characteristics. To the extent that no two patients can possibly have the exact same set of characteristics and life experiences which may be considered a priori as relevant to their psychopathologies, no two such populations can be said to be identical. However, in the process of learning from experience, particularly in deciding how to treat a given patient based on previous experience with other patients, we do not demand that our given patient's relevant variables all be identical to those of the earlier patients whose treatment we are using as precedent. The decisions are necessarily based on similarities rather than on complete identities. That is, with respect to the single-case population under discussion, the application of the results of the sampling of such a population to other single-case populations must be performed in terms of similarities of relevant characteristics. From the point of view of purposeful clinical research design, the systematic application of the stochastic model to increasing numbers of patients carefully selected for similarities and dissimilarities in particular patient characteristics then becomes the basic statistical approach toward identifying the factors most relevant to the success of treatment, as well as in attacking the problem of the more general evaluation of treatment itself.

While the essential purpose of this presentation is to stress the intensive statistical study of the single case as the basis of clinical design, a single case is certainly not the optimal number for every piece of clinical research. All things equal, many intensively-studied cases are better than one, and a great many intensively observed are better than just many. Within a given set of resources, however, the more intensively each patient is observed, the fewer are the number of patients that can be observed. In a statistical study, with from four to eight patients observed intensively over time, one can, for example, divide the patients into two groups, to facilitate control along the time axis, and to complement in this way the use of each patient as his own control.

Some comments on the elasticity of the stochastic model

Designs based on the intensive statistical study of each case also allow an elasticity which is impossible within the extensive model. The accumulation of data, say, on a daily basis, allows an ongoing statistical analysis of results during the course of a study, which can suggest useful modifications of the design as the study progresses. Within the framework of intensive design, one can also take into account and analyze, over sufficiently long periods, the effect of various day-to-day contingencies and inter-

current events such as parental visits, ward altercations, and so on. Stochastic designs in which a number of variables are observed on a day-to-day basis with respect to each patient can also contribute useful data concerning the timing of various changes in the patient-state and the correlation of changes over time within a patient with respect to various aspects of his psychopathology. Similarly, day-to-day observations on each of a group of patients allow an approach to the analysis of contagious aspects of psychopathology and interpersonal processes by the correlation of data among patients along the time axis.

Many behavioral and affective variables relevant to the study of psychopathology have in recent years been recognized as having important interactive components, and in some theoretical frameworks the interpersonal aspects of psychopathology are regarded as basic. In any event, the day-to-day influence of the observer on the ward upon the behavior and symptoms of the patients is generally recognized. Variations among observers concerning the evaluation of a patient with respect to those characteristics of psychopathology which have interpersonal components are then not simply a reflection of differences in the evaluation of the patient, but are also a product of differences in the relevant aspects of the interpersonal fields generated between the patient and his respective participant observers.

A stochastic design in which a number of observers evaluate the same patient over time affords an analysis of some aspects of interpersonal phenomena as part of the evaluation of the patient-state and its possible transformation following the introduction of a new therapy. Such an approach further allows an analysis of differences in patient evaluation, which may be a function of interobserver differences in personality and previous relevant interpersonal experience.

As an example, let us suppose that a patient has two observers who evaluate him independently each day on some behavioral rating scale or check list. Let us assume further that one of the observers is known to be permissive, while the other is recognized as having a somewhat authoritarian attitude. It is not unlikely that these personality distinctions may result in different patient-state interpretations with respect to a number of variables which commonly appear in various clinical evaluation or rating forms. Such evaluation differences may be the result of differences either in interpersonal fields, or in past interpersonal experience resulting in corresponding differences in expectation concerning the patient, or, quite likely, both.

Now let us assume an intensive study of the patient over two successive time intervals, such that during the second of the intervals a new treatment factor is introduced, and suppose the following results are obtained with respect to an evaluation of "hostility" (presumably the patient's hostility): The reports of the permissive observer average to a state of minimal hostility during both the first and second period; the report of the authoritarian observer averages to a marked hostility during the first period, and a minimal hostility over the second. That is, as far as the first observer was concerned, there was no essential change in the patient's degree of hostility between the two periods, while to the second observer, an apparent improvement has taken place from a markedly hostile to a minimally hostile state. Does the fact of this disagreement diminish the importance or the usefulness of these data? Is it even necessarily a question of which observer was correct? The answer to both of these questions is no. It is useful to be able to estimate that as a result of a new treatment a patient appears less hostile to an authoritarian figure, even though a more permissive person may have estimated a minimum of hostility in the patient from the start. The hypothesized set of data can thus unambiguously be stated as representing some degree of change in the patient.

The essential argument can be stated in more general terms. Insofar as psychopathology may be considered as manifest within the framework of interpersonal events, the extent of an individual's psychopathology may be considered as a resultant of the distribution of interpersonal

fields of which he is a member and which he helps to generate. An approach toward an objective evaluation of at least certain aspects of a person's degree of mental illness or mental health therefore requires some estimate of the distribution of feelings and attitudes he arouses in others. As interaction with the patient will to an extent depend upon the observer's own personality and set of attitudes, some knowledge of his relevant characteristics can obviously provide more useful results concerning the patient-state and its transformations than if all observers were considered as identically calibrated for patient observation and evaluation, and their individual traits entirely ignored.

Increasing reliability in the stochastic model

There was no intent in the discussion of the previous section to deny the need for consensual validation or interobserver agreement concerning the use of such devices as rating scales, check-lists and so forth. Much of this aspect of the problem of reliability has to do with the question of agreement on the way words are to be used in describing various manifestations of the patient-state. One great handicap to achieving reliability in this fundamental sense is the existence of the great variety of outward manifestations of many dimensions of behavior which can be observed from patient to patient. For example, there are many ways in which a patient can be hyperactive or bizarre. The specific observable phenomena to which either of such relatively abstract words can be applied are obviously quite numerous.

If one is oriented toward the extensive model of design in which the basic statistical tests and comparisons are made between groups of patients, it seems natural to develop rating forms, check-lists, etc., whose terms are sufficiently general to include the various examples of a single dimension or broad aspect of behavior under a single word or term, or at best, under a few such terms for each dimension. (If too large a number of terms are used, then many of them will not be applicable or relevant to the description of a sizable proportion of the patients under study.) It is this use of high-level abstractions which contributes heavily to the difficulties of achieving agreement in the description and evaluation of the patient-state. Once the emphasis is placed on the intensive statistical study of the single case, it is a natural development to think of forms, check-lists, etc., designed especially for the individual case. One can then stay close to the specific manifestations of a given patient's behavior and affect and other possible symptomatology in the construction of such forms. For example, it would be much easier for two observers to agree on whether or not on a given day a patient wears his shirt inside out than whether he is "bizarre." In order to use such an approach, one has, of course, to know the patient well. A preliminary period of observation, as well as a review of the patient's case history including nurses' notes and so on, will generally be helpful in devising a form for the systematic observation of the patient within the stochastic model, and which will be particularly relevant for the patient selected for study.

The suggestion that reliability can be increased by the use of reporting systems tailored to the individual patient's psychopathology is not made at the expense of any real loss of comparability of results between patients. On the contrary, the increase in reliability which the approach affords will increase the validity of such comparisons. Thus, while statements about increases or decreases in hyperactivity can be made with greater accuracy once a patient's hyperactivity is first reported in terms of the specific ways in which the hyperactivity is manifest, the more general or abstract term can be brought in for purposes of comparison and generalization *after* the more specifically defined data have been collected.

Clinical responsibility and research interest in the intensive design

It is apparent that an approach to clinical research which stresses the detailed study of the individual case is bound to stimulate interest in research in many clinicians who have felt remote, or at best, uncomfortable about participation in research projects based on the extensive model. The fact that

the stochastic model can validly be applied to a small number of cases, and with sufficient care even to a single case, allows the clinician to engage in meaningful research projects to satisfy some of his own interests without the need for the elaborate and expensive controls of the extensive design.

REFERENCES

Chassan, J. B. On the unreliability of reliability and some other consequences of a probabilistic definition of the patient-state. *Psychiatry*, 1957, 20, 163–171.

Chassan, J. B. On the development of clinical statistical systems for psychiatry. *Biometrics*, 1959, 15, No. 3.

Kety, S. S. Biochemical theories of schizophrenia, I. *Science*, 1959, Vol. 129, No. 3362, 1528–1532.

Kramer, M., & Greenhouse, S. W. Determination of sample size and selection of cases. In J. O. Cole and R. W. Gerard (Eds.), *Psychopharmacology: problems in evaluation*. NAS-NRC Publication No. 583, Washington, D.C., 1959.

(Manuscript received August 26, 1959.)

Can you remember how acute your sensations were, how intensely you felt about everything, when you were a child? The rapture of raspberries and cream, the horror of fish, the hell of castor oil! The inexpressible joy of sitting next to the driver, with the smell of horse sweat and leather in one's nostrils, the white road stretching away to infinity, and the fields of corn and cabbages slowly turning, as the buggy rolled past, slowly opening and shutting like enormous fans! When you're a child, your mind is a kind of saturated solution of feeling, a suspension of all the thrills—but in a latent state, in a condition of indeterminancy. Sometimes it's external circumstances that act as the crystallizing agent, sometimes it's your own imagination. You want some special kind of thrill, and you deliberately work away at yourself until you get it—a bright pink crystal of pleasure, for example, a green or bruise-colored lump of fear; for fear, of course, is a thrill like any other; fear is a hideous kind of fun.

ALDOUS HUXLEY, *The Genius and the Goddess*

DEDICATION OF THE MENTAL HEALTH RESEARCH INSTITUTE BUILDING

Dedication ceremonies for the new Mental Health Research Institute building at The University of Michigan were held in Ann Arbor, Michigan, on January 29, 1960. Following are excerpts from talks that were presented during the day's ceremonies.

MORNING SESSION

Dr. Harlan Hatcher, President of The University of Michigan, presiding.

President Hatcher: It is a moment of pleasure for me to open this session which will be both a dedication of a new, important facility for the University and a professional conference on the work that lies behind this effort in the interest of man's fundamental welfare. It is a significant occasion for us in many respects. It has great personal significance to me as it has to many of the people who are gathered here including, particularly, Dr. Albert C. Furstenberg (former Dean of the Medical School), Dr. Raymond W. Waggoner (Chairman of the Department of Psychiatry), and Vice-President Marvin L. Niehuss (of The University of Michigan). I single out these men only because I am certain that as we center our attention on this dedication ceremony today our thoughts go back some four or five years to the initial discussions, the long and serious periods of planning, the doubts that were raised, and the optimistic feelings we had that this was really a venture worthy of the finest efforts of all of us. The founders of the Institute came here under circumstances that required faith and optimism because it was something of a problem, even four or five years ago, to find some cubbyhole or corner where they might have a headquarters to begin their work. But they did come, started their work, and now these few years later we are able to meet to dedicate a facility to their use and to rejoice with them in the scientific progress which has been made since they began their labors here.

We have brought here a group of men and women who are joined in a broad campaign to attack mental illness at its very base. We are confident that these researchers who are delving simultaneously into various phases of the problem will in time devise ways to help man live more at peace with himself and with his environment. The statistics of mental illness, so familiar to everyone in this room, come as a shock to most laymen: nearly three-quarters of a million people hospitalized in mental institutions, millions more under private care, an annual bill of some three

Invocation

O thou who are the life and sustainer of all, we would remember first at this time that company who in every age have given themselves to science, from the inventor of the wheel to whoever discovered or put something to work this morning.

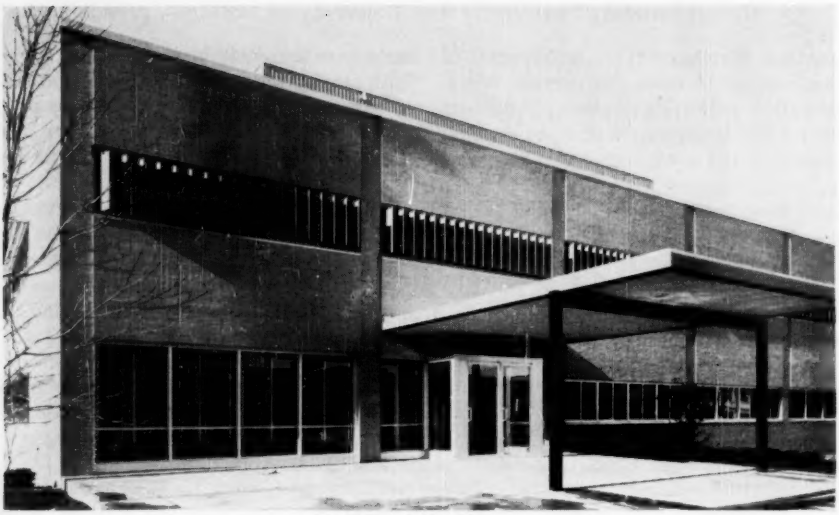
We thank thee, that in their steps, by research and invention, we share in thy creation, in the faith of thy word that this world is all very good, and that body and mind are both alike to thee.

May the Mental Health Research Institute ever be useful. Bless its members through the years, and reward the leaders of this commonwealth and the others who have made its work possible.

Keep ever before us the needs and anxieties, the handicaps and burdens of thy people everywhere, that these also may stir us to greater devotion to science.

Help us all to live nobly, to thine honor and glory, and to the good cheer of our fellow men, and to thine Incarnate Word be the praise as it was in the beginning when the morning stars sang together and all the Sons of God shouted for joy.

The Reverend C. Leslie Glenn



The Mental Health Research Institute Building

billion dollars, an annual loss of 325,000 man years—this is a great toll. We here at The University of Michigan are grateful to the people of this state for giving us the opportunity to help combat this monumental problem. We reflect once again that The University of Michigan has indeed been fortunate throughout its history in having the enlightened support of the citizens of this state. The contribution of the United States Public Health Service toward this fine new building in addition to the great support they have given in the programming is likewise acknowledged by us with deep gratitude. The McGregor Fund of Detroit deserves and receives our profound thanks for the splendid equipment that was made possible by its donation.

This event also symbolizes the University's determination to press forward in all fields of knowledge, not to specialize in just a few areas. As our attention is drawn irresistibly to the planets, our galaxy, and the galaxies beyond, it is reassuring and essential to find here a body of scientists concerned with the innermost recesses of man's mind. There is a certain irony in the fact that the same brain which has enabled man to rise to his feet and strain for the stars also causes him so much personal torment.

These occasions, as I have said, are happy ones. They represent solid accomplishments and they inspire some optimism for the future. I like to think of the new concept, still fragmented in the minds of scores of scientists, which will take shape in this Mental Health Research Institute building, and of those ideas which will ultimately come to benefit millions of our people. The challenge, opportunity, and satisfaction lie ahead of us. I share your pleasure in gaining a superb research facility and wish all of you who are concerned with it the best of success.

We are honored this morning to have as our special guest the Governor of our state, G. Mennen Williams. His long-continued interest and concern for all phases of education and research are well known to every citizen here. This morning we would take special note, however, of his very deep and dedicated concern over the years for the field of mental health. From the very beginning of the idea that The University of Michigan would be the home of this particular research unit, Governor Williams has not only been involved but has given it his strong, whole-hearted support. Governor Williams, it is a special privilege for me this morning to present you to this conference and to call upon you for your words, both for yourself and for the State of Michigan.

Governor Williams: It is a great pleasure to join with you today at the dedication ceremonies of the Mental Health Research Institute Building of The University of Michigan.

This is an unique institution and it is fitting that Michigan should be its home. We are proud of Michigan's improvements in the mental health area over the last decade. We have written a nationally acclaimed record. Permit me to mention only a few of our tangible achievements: We have added 6,600 beds to our state mental health system, an increase of 23⅓%. We have ten new child guidance clinics and five new adult guidance clinics. These treat approximately 360% more people each year than the clinics of a decade ago. Outpatient facilities, previously non-existent, now service over 2,000 patients a year. Discharges due to recovery have increased 28%. While our population has increased 27%, our admissions to mental hospitals have gone up only 18%. Patient discharges today roughly balance admissions. And naturally we are proud that our new Northville State Hospital was one of only 16 hospitals in the nation to receive complete approval by the American Psychiatric Association's Inspection Board. Finally, I must mention our specialized facilities: the Children's Psychiatric Hospital here in Ann Arbor, the Lafayette Clinic in Detroit and the Hawthorne Center near Northville, each unquestionably one of the nation's outstanding.

It is impossible to measure these things in dollars and cents, but it has been estimated that without our improved prevention and cure program we would have an additional need of $23,000,000 worth of beds, over and above what we need now.

With this background it is perhaps inevitable that Michigan should pioneer in the mental health area with this research institute. The need for an institute of basic research in mental health has never been greater.

Now I know it is possible to stack the statistical deck and prove almost anything. But the mental health field is one field where, unfortunately, no matter whose statistics you take, the picture is far from rosy. And this is true despite the achievements of states such as Michigan.

At mid-century, 8,500,000 Americans, or one out of every 18, was suffering from some sort of mental illness. In the course of his life, one out of every ten Americans will need some kind of mental care. Nationally, there are as many patients in mental hospitals as in all other hospitals combined. An equal number of mental patients requiring hospitalization are waiting their turn for admission. Let me say that Michigan's hospitals at Pontiac, Traverse City, and Ypsilanti were given only conditional approval by the American Psychiatric Association Inspection Board solely because of forced overcrowding. Close to 1,000,000 Americans who reported for the military draft were found to be suffering from some kind of psychic difficulty.

Confronted with these statistics, and many more, it becomes apparent that an institute of mental health research is not a luxury. It is a necessity. It is a necessity for the economic health of the state and the nation as well as the mental health of the people of this country and the world. In these troubled times, this building will stand as a symbol of sanity. I do not mean it is a symbol of sanity because it deals with mental health problems. Rather, I suggest that it is a symbol of sanity because it stands for the importance of basic, fundamental research at a time when people have difficulty seeing beyond the tailfins on their cars and the coming color on their TV sets. Further, it stands for research into possibilities of *preventing* mental illness, not merely diagnosing and curing it. Today such research is imperative. Our future stands imperiled by our ignorance. The place to start developing healthy minds and personalities is not in our mental hospitals.

While we rejoice at the establishment of this new building, we must remember that it has been long overdue. For some time now Americans have been exposed to key psychoanalytic concepts. In the social sciences, psychology rivals economics as a relevant factor in explaining current and past events. In the movies, the theater, the novel, we see hidden memories unveiled and psychological blocks removed. Paperback books parade the works of Freud and Jung. Even in the world of politics I find that occasionally instead of accusations of corruption, one politician will accuse another of suffering from some sort of complex.

Despite the currency of psychological concepts and jargon, the people of this country have been reluctant to proceed, full speed ahead, with a series of fundamental psychological research programs to further explore the great unknowns.

Perhaps a major reason for the failure to move ahead in this difficult area is that Americans habitually demand quick results, and research into the possibilities of prevention is no overnight task. We like to push a button, turn a key, press on a pedal, pull a lever—and then watch the results. Indeed, the museums of science and industry, where the flick of one switch sets off a Rube Goldberg galaxy of gadgets, capture our naive conception of science. We have been willing to buy now and pay later. But we have been unwilling to pay now and reap the rewards later. It has taken Americans a long time to realize that research takes time.

Despite our tendency to look for quick relief, magic formulas, and wonder drugs, I think there is a growing awareness of the need for basic and advanced attacks on the national mental health problem. Perhaps the conjunction of juvenile delinquency, quiz show and payola scandals, and the jolt of Soviet scientific achievements have caused us to ask ourselves some important questions about our personal and public values. Indeed, some believe that America itself today suffers from a national schizophrenia,

that we are unable to connect public and private values.

In any event, I know that we in Michigan have a growing awareness of the importance of fundamental research to the future. Our new Institute of Science and Technology, together with this building, augurs well for the years ahead. And Michigan's faith in our Institutes has not been a lonely one. For instance, the Mental Health Research Institute has received grants for specific work from the U.S. Armed Forces, the U.S. Public Health Service, the National Research Council, and various private foundations interested in furthering basic research in mental health.

Perhaps the measure of Michigan's achievement is our realization that the struggle has just begun.

President Hatcher: Thank you, Governor. I should like to recognize the Director of our Mental Health Research Institute, Dr. James G. Miller, and give him the honor of presenting our next speaker.

Dr. Miller: As Governor Williams has pointed out, there have been contributions from the state of Michigan, from the federal government, and from private sources toward the construction and equipment of this building. We felt it appropriate, therefore, to ask a representative of the United States Public Health Service, the agency which contributed the federal funds, to speak this morning. Dr. Jack Masur, the Assistant Surgeon General of the United States, was to have been with us but he is just recovering from the flu. However, he sent as his representative Dr. Harold Hildreth of the United States Public Health Service, who came through snow and sleet and rain and fog in order to be with us today. Dr. Hildreth is a psychologist, formerly Chief Clinical Psychologist of the United States Veterans Administration, now working in the Community Services Branch of the Public Health Service. Dr. Harold Hildreth, speaking for Dr. Masur.

Dr. Hildreth: Representing the United States Public Health Service and the Department of Health, Education, and Welfare on this important occasion is not the Assistant Surgeon General who looked forward so eagerly to being with you; nor, for that matter, any of several others who might more appropriately than I have appeared in his stead. But because of the exigencies of weather and plane flights, sudden work, urgent work in a common cause, overwork to the point of illness, there stands before you the last of the Mohicans on this particular day. But do not conclude, please, that it is only by virtue of indolence that I am here. I planned early on coming in the sweet anticipation of a whole day's luxurious listening. A day, as the radio puts it, planned for my listening pleasure. Dr. Masur, whose distress was acute at having to cancel his trip here at the last minute, asked me to read to you his prepared statement which reveals, as you will see, the far-reaching significance he attaches to this occasion.

Dr. Masur: I am proud to have been asked to join you in dedicating this splendid new facility of a great University. The formal opening of this special laboratory seems particularly fitting as a new decade begins, because the advent of the 1960's marks, in many people's minds, a time for looking ahead and for seeking out new directions for progress.

This University, one of the proudest traditions of the people of Michigan, is concerned with the acquisition of new knowledge, the dissemination of accumulated knowledge, and the application of all knowledge. As has often been pointed out, these roles of research, teaching, and service are closely intertwined. Research enhances the vitality of teaching; teaching lifts the standards of service; service opens new avenues of investigation.

The fields of mental health and mental illness are only belatedly receiving the attention they deserve in the structure of research activities in most of the centers of learning in this country. The almost overwhelming burden of "service" requirements for the millions of emotionally sick people has left too few men and women available to engage in integrated clinical and laboratory investigation of mental disease. The paucity of trained, qualified scientists ready, able, and willing to enter this most

complex of all health problem areas has retarded the advancement of knowledge.

The gathering together of a group of eminent workers, the environment of a renowned University, the construction of a suitable facility, the assurance of continuity of support in the scholarly pursuit of knowledge—all of these add up to why this is an event of special significance in the health of the Nation. You have added a new and critically needed center to the landscape of American research into nontechnological fields. The work being carried out here is of urgent human importance and will assume even greater rank as the years go on. The potential for human betterment through the knowledge your staff seeks, and the clinical and laboratory investigators you train, is great—very great indeed.

As a representative of the Public Health Service and the Department of Health, Education, and Welfare, with their responsibility for the stewardship of public funds for medical research, I want to stress our role of partnership—a persistent partnership as the President has expressed it. We believe that the Government's role is one of appropriate sharing of burdens and responsibilities and leadership. We believe that the creation and sustenance of the Mental Health Research Institute offers a fruitful opportunity for a collaborative undertaking of notable importance. All of my colleagues join in wishing for a long, productive, persistent partnership. In this Institute you have embodied the true ideal of the University in bringing together scholars in search of knowledge useful to mankind. Let me borrow the words of Thomas Huxley:

> "So sure is it that men live not by bread, but by ideas, so sure is it that the future of the world lies in the hands of those who are able to carry the interpretation of nature a step further than their predecessors;
>
> "So certain is it that the highest function of a university is to seek out these men, cherish them, and give their ability to serve their kind full play."

President Hatcher: Thank you very much, Dr. Hildreth, we appreciate your coming. I hope you will convey to Dr. Masur our regret that he was unable to be with us, our hopes for his speedy recovery, and our special thanks to him and all of his colleagues for the continuing interest and aid which they are giving to us and to all units of this type.

I have three very special reasons for taking unusual pleasure in presenting Dr. Ralph W. Tyler to this group now. The first one is entirely personal because he and I are friends, both personal and professional, of long standing, dating back to the days when we were young men together at Ohio State University. Secondly, it was Dean Ralph Tyler who exerted the imaginative administrative leadership at the University of Chicago which was importantly instrumental in enabling this distinguished group of scientists first to come together for their joint attack upon this area. Third, of course, is that Dr. Tyler, Director of the Center for Advanced Study in the Behavioral Sciences, is a national figure, one of the outstanding leaders who have kept before us always the tremendous importance of the whole field of the behavioral sciences in a period when we have been quite properly so elaborately intrigued with the physical sciences. We are honored by his presence and I now take great pleasure in presenting him to you.

Dr. Tyler: I want, first, in representing the so-called behavioral sciences, or the sciences that are concerned with the study of human behavior, to follow what previous speakers have done: to praise the state of Michigan, the United States Public Health Service, and the McGregor Fund for providing the funds to build and equip this fine structure. Those of you who have not seen it and will have a chance to do so later in the morning will be greatly impressed with what a beautiful and adequate structure it is for housing the Mental Health Research Institute. This is something of which the state of Michigan can be very proud. There is nothing like it anywhere else.

I want, second, to emphasize the importance of research in mental health, although I do not need to dwell upon that because both the President of the University and the Governor have spoken of the importance of the problem we face: the

need for much more adequate understanding of the nature of mental illness, of mental health on the positive side, of the factors that condition and bring about the deterioration of mental health, and of the means of prevention or at least of a reduction in its incidence. There never was a time when the need for research of this sort was felt more greatly by the American people and by people throughout the world, for this is not just an American phenomenon.

And now I want to emphasize, because I think this is the genius of the new Mental Health Research Institute at The University of Michigan, the recognition that mental health and mental illness are very complex. They are not simple diagnoses. It is not like diagnosing a particular disease of the body in which you can identify the particular place where something goes wrong. The term 'mental illness' refers to the failure of a human being to function in his life. Such failure involves a multitude of factors within his environment. The person who may be able to function quite successfully in one situation may break down in another—which calls attention to the fact that to understand mental health is to understand much more about the nature of our total society, of the environment in which we live and the demands it makes on us, of the kinds of defenses we build against difficulties, of the ways we can retreat from society when difficulties get too great, and of the defects in our own personality as well as in our physical systems.

Mental health and mental illness represent the disease not alone of the individual but the disease of society. It requires for its solution an attack that is broader and deeper than the attacks on perhaps any other kinds of illnesses that we know about, although increasingly we are recognizing that many other illnesses have their particular problems in relation to the nature of the environment in which we live. But because of this great complexity, because of the need for embarking on research on a scale that will look at the total range of factors, our greatest hopes for research lie in the kind of thing that the Mental Health Research Institute stands for. Here they are dealing with the individual and his environment in all the complexities of both, then working on general theory that will provide concepts and principles that have wide application and can bring order out of what would otherwise be an impossible complexity. These are key points in the development of the Mental Health Research Institute, as I understand them from reading the reports and from long conversations with its members. This, it seems to me, is a great goal, and a very difficult goal, for which The University of Michigan and the state have pledged their support for the building and have made it possible for these people to be here.

President Hatcher referred to the fact that the group now comprising the Mental Health Research Institute has a history. I am happy to have had some part in that history. It was more than ten years ago when Dr. Miller was chairman of the Psychology Department at the University of Chicago and I was Dean there, and a number of other persons who have been associated with the Institute here including Drs. Gerard, Rapoport, Crane, Easton, and Marschak, became actively interested in the possibility of this kind of understanding and attack upon human behavior. They came to me to get support and encouragement and to issue the invitation to bring together a working group in what they called the Behavioral Science Study Group. This group was at work for several years with difficulties and with successes. Difficulties because in the first place there was no adequate pattern to guide the effort to look at such a complex phenomenon and to do so in a way that brings in people from a variety of academic fields, including neurophysiology, biochemistry, other medical and biological fields, the social sciences (including such diverse disciplines as political science, economics, history, and psychology), and the physical sciences, particularly mathematics. This was not an easy task.

It was difficult, too, because of the fact that people have not been trained to relate their disciplines and their ideas to other areas. It has been important, of course, to

have highly specialized types of attack upon problems, to break down a problem until it is simple enough that it can be investigated without confusion from a wide variety of factors. Much of science has developed through this degree of specialization and through increased fragmentation. But science—particularly the physical and biological sciences in advance of the social sciences—has also developed through synthesis, through beginning to see common theory and common measures. For example, physics and chemistry have many things in common in the atomic structure of matter, in the notions of the relation between energy and matter, and the relations between chemical and physical change. There is in a sense a stimulation from other branches of science, and yet there are real problems and difficulties in implementing synthesis. But this group moved on and made some progress though they found it difficult to find people who felt competent to carry on in this fashion.

In 1955, the opportunity to come to The University of Michigan enabled the group to move here and to add both to its staff and to its facilities. Shortly, too, after coming to Michigan, they founded the journal *Behavioral Science* which provides a clearing house for research reports and ideas related to this interdisciplinary attack upon human behavior, human problems, and human understanding.

The development of the research program during the past five years has been to me one of the most impressive aspects of the Institute. Among the problems that arise in interdisciplinary thinking is the tendency for people to take a large theory and without making any empirical tests of it carry it over to other problems, so that it becomes really a large, informal category. Things *look* like similarities in nature but they look that way partly because the categories are so big. When they are given more concrete embodiment and data, when they are actually examined more carefully and measured and described more accurately, many of the similarities disappear. But this group has been able to bridge that gap. It is consistently testing its theory in empirical data, seeking to find out from actual experiments the extent to which these large theories are useful and what necessary conditions and modifications are required to make them apply to human behavior, to human functioning, and human breakdown.

Another important problem of interdisciplinary research is the fact that, typically, concepts used by scholars in one discipline are not understood by those in another discipline. When such workers try to put things together, they are like the tower of Babel, unable to speak the same terms. This group has had over the years continuing seminars and study conferences so that they now talk a common language—they understand what their terms mean so that economists can speak the language of psychologists and biologists.

Finally, a third problem, one that I think has faced every interdisciplinary effort in the sciences, has been the problem of developing new scientists, scientists for the future who occupy what are sometimes called interstitial roles; that is, they represent people who know several areas of study and have these as part of their whole system of science. A case in point is biochemistry. It is now common to have people receive Ph.D.'s in biochemistry. They are trained in what were originally aspects of physics, aspects of chemistry, and aspects of biology, but they have added many new phenomena which are obtained when you begin to examine what is called biochemistry, the nature of the chemistry that goes on in living organisms and living processes. People then become able to carry on the concepts that the particular techniques require to handle a new discipline. Such training is necessary in the Mental Health Research Institute, and I understand that as part of the educational or training program, they are beginning to develop some of these interstitial or interdisciplinary people who have in their own background and training the kinds of techniques and the kinds of concepts able to handle the relationships among some of the disparate disciplines. This, it seems to me, leads us to have confidence that the kind of attack here,

although extremely difficult because human behavior is as complex as anything we know, has great probability of moving forward successfully and of becoming a model which will be followed in many other places.

In conclusion, I want to express from the Center for Advanced Study in the Behavioral Sciences, where we are concerned with the problem of broader and more adequate understanding of human behavior, the hopes that this Institute will be a pioneer and continuing leader in the development of new knowledge and understanding of mental health, of mental illness, and of human behavior generally. The building, I think, should facilitate it. We have doubters, like Parkinson who maintains that any organization reaches its capacity and its greatest effectiveness just before it enters into a new building and that thereafter you can expect no great effectiveness from it. On the other hand, I would prefer to end with the speech that was given by Robert Maynard Hutchins at the dedication of the Social Science Research building of the University of Chicago 30 years ago last month, a building, by the way, which facilitated a great deal of the kind of research activities we are talking about within the social science field. Dr. Hutchins said, "This building means that social scientists at Chicago have an opportunity and an obligation. If co-operative research is possible at all on a large scale, it should be possible here and now. If social research has value, it should be revealed under conditions so favorable as these. If the type of organization of the social sciences here is useful, that should be made plain, for no physical difficulties stand in the way of success. If this building does not promote a better understanding of our society, we shall know that there is something wrong with the social sciences or something wrong with us, for here for the first time everything that can serve the social investigator is ready to his hand. If he cannot make progress now, he never can. The building, then, is in the nature of an experiment, a test case, to determine whether with such facilities and such an atmosphere we may more rapidly work out the problems that confront us." I think this applies to the building of the Mental Health Research Institute.

President Hatcher: Thank you very much, Dr. Tyler, for being with us and for your presentation to us. It is now my very deep personal pleasure to present the distingushed speaker who will conclude the morning session. He is one of Michigan's own beloved sons and we have come to feel that the only proper introduction for Dr. Detlev W. Bronk is just to welcome him back home. It would be idle for me to outline to you the remarkable achievements of this great scholar and educational leader, his many, many services to his government and to his own chosen field of science, his period of leadership at Johns Hopkins University, and now his truly great work as President of the Rockefeller Institute, his accomplishments in the whole field of science and especially in medicine, and his service as president of the National Academy of Sciences. Dr. Detlev W. Bronk.

Dr. Bronk: This being the third building at The University of Michigan at which I have been privileged to participate in the dedication, I am beginning to feel that I almost deserve the right. While I have high regard for Parkinson's Law, I do not hold with him that man works without the means to work. I have recently been interested in the movement of the field of oceanography. When the National Science Foundation a few months ago made a grant of $3,000,000 to the Woods Hole Oceanographic Institution, someone asked why this was necessary. He remembered that Henry Bigelow used to do his research in the salty bottom of a fisherman's dory, but I do not find that a fisherman's dory is adequate for the exploration of the seas. I do not believe that odd attics and empty halls make research possible.

It is embarrassing to one with some modesty to force one's deliberations upon another, unless one has, as we in Quaker meetings say, a concern. I have some concerns which I cannot say are significant at all, but I hope they may prove somewhat relevant to the occasion and the times.

I first would speak on this because you

are dedicating a new facility which has been provided by generous government and noble individuals. And my concern in this. Why is it that we of the United States of America are today largely supporting science? If you take a reasonable estimate, I would say that we are spending approximately ten billion dollars on science for this current year, which includes both research and immediately related developmental undertakings. It has been my observation that one of the principle reasons why so much support is given to science is because of fear. I hasten on to say that much of this fear I believe to be reasonable fear. I remember hearing Frank Aydelotte, who was President when I was Dean at Swarthmore College, say at one meeting that to have too much fear paralyzed one's efforts, but to have no fear meant to folly. I would say that the fear of what antagonistic states may do is a very good reason for the expenditure of large sums of money in support of science. Another reason why we spend so much is because of fear for ourselves and our loved ones that we may become ill or die. This, I would also say, is a reasonable fear and is good justification for support of science. But then I go on to say, what is it that we wish to have in the life which we will defend with our armed might and in the life which we will prolong through medical science? I am reminded of a phrase I may myself have used regarding one I knew and all I could think to say on the occasion of his death was this: Throughout the long years between his birth and death he lived so little. I take it that what we want to do in defending our way of life is to enable us to live—to live more richly and more fully. I would like to state that many of us have supported science because it is more than a means of defending our borders, our way of life, and life itself; because, in fact, science is a great mental and spiritual endeavor. I would wish that more would consciously support science because it is recognized as an essential element in the evolution of man.

Vast changes lie ahead, and this gives me another concern of which I would briefly speak. If you ask me why it is I think vast changes occurring at an accelerated rate lie just ahead, I would say what we all know. The people of our shrunken world have learned what can be had from science and technology. Those of us who are concerned with the national program for foreign aid in its great diversity of efforts recognize that we are satisfying the desires of people who are living in an era of expanding expectations. This will, whether we like it or not, accelerate scientific effort, cause us to support it more generously with financial aid, cause us to train more men and women for careers in science and support them more fully. It will cause us, I believe, to educate more people to understand what science is. That is one reason why vast accelerated changes in our way of life lie in the not-far-distant future. Furthermore, because of the great emphasis upon military technology, there are bound to be numerous by-products which will be applicable not only to our armed forces but to all in all walks of life. Finally, there is the new spectacle of nations competing in scientific achievement because it is believed that it brings national prestige around the international conference table and in the battle for the loyalties and minds of peoples throughout the world.

The past 25 years have shown dramatically how profoundly we can change and shape man's physical and social environment. Transportation is a natural example of this. I need not say to an audience in this state of Michigan that the whole system of highways which have been made necessary by the automobiles you have given to the people have transformed our way of life. I need not say to anyone who lives in an urban environment that the pattern of living is completely different and is rapidly changing from what it was a quarter century ago. We have changed our way of life profoundly. We have given man great new powers of speak and motion, of human effort. We're dealing with a vastly growing body of knowledge. My concern is this: what do we choose to do with these new powers?

There was a fellow student at The University of Michigan when I was here who, a few weeks ago, had an editorial which was reprinted by the Westinghouse

Television Corporation in the *New York Times* decrying the influence of the intellectual. I took down my Skeat's Etymological Dictionary to reassure myself that I knew what an intellectual is. Although I am not an etymologist, if I understand it correctly, an intellectual is one who is trained in making decisions—it means to choose between. What is wrong, I would ask, with emphasis upon the training of the mind so that people can choose between this and that? One of the great challenges of every university is to train people to choose. You can take a motor car on a Sunday and drive 200 miles and see, as Governor Williams said, the tail fins on the car ahead. You can see cluttered highways which have despoiled a lovely countryside, or you can use the car to go out into beautiful nature to see friends with whom you have stimulating converse. You can use television for the exchange of great ideas or for the marketing of cheap and sordid notions. You can take the great accomplishments of a city—you can use it to the glory of mankind or you can make it a sordid and a miserable place in which people must live unhappily together.

I have been speaking of these concerns because I believe that we not only will profoundly change our way of life within the next few years, but I also believe that we can use the developments and the applications of science to create a more desirable way of life. But then I must ask, do we know enough about man so that we can create this more desirable way of life?

The problem which Dr. Miller, Dr. Gerard, and their colleagues have addressed themselves to is not easy. Because it is not easy, I believe it is all the more important. And this is my last concern: that we support with enthusiasm and stakes the efforts of men and women who are determined to try to understand man and his behavior—to try to understand why it is that some people will not give to others what they demand for themselves, why it is that people who respect their own dignity will not respect the dignity of others, why it is that some are unwilling to preserve for the future of their descendents the good things they crave for themselves now. I could go on through an endless list of queries that every thoughtful man poses to himself. I repeat that science gives us the building stones of a better world but the world will be as we choose to make it. Unless we have more people such as those working in this new research institute who endeavor to study the characteristics of behavior—to understand the nature of man to all its fullness—the world will not be what it could be.

I would like to speak once more about transportation. I was coming back from Washington a few months ago and I had a few experiences I would like to recount. I passed the American Airlines desk and there was an angry group of men and women—angry because the plane for Los Angeles was delayed one-half an hour. As I went to make a phone call, I heard an angry man berate the operator because she said there would be a five-minute delay on his telephone call to Kansas City. I wondered, what is it in human behavior, what is it in the spirit of man that makes him unwilling to appreciate and recognize the vast powers which have come to him through the evolution of his mind—or call it science and all other creative efforts if you will. I flew into New York. It was a glorious, moonlit, starlit evening and as I saw my native city spread out before me in a multitude of lights, when I reflected on the fact that in 55 minutes I had come that long distance from our nation's capital to the metropolis, I felt a surging cry in being a man who has made it possible for men to do such things as that. Then I drove back to our home in the Carlisle Towers through some of the most sordid slums man has ever built and my pride went back to normal size because I saw the great gap between what we can do and what we do. I was reminded of the lines of Edwin Markham, I think it was, who said,

> "Why build these cities glorious if man unbuilded goes?
> In vain we build the world unless the builder also grows."

That understanding is the challenge before this new institute.

LUNCHEON MEETING

Dr. Raymond W. Waggoner, Chairman of the Department of Psychiatry of the University of Michigan Medical School, presiding.

Dr. Waggoner: During the past 20-odd years that I have been Chairman of the Department of Psychiatry at the University of Michigan Medical School, we have received strong support from the Legislature of the state of Michigan in bringing about excellent facilities for teaching, training, research, and clinical service in the field of mental disease and health. For this expression of support and confidence, I would like to express the appreciation of the Department of Psychiatry and of myself, and to pledge that we will make the best possible use of the facilities thus provided.

Some years ago it appeared that it would be of value to develop a program for research in the field of mental disease which would not be concerned primarily with project research but rather with basic research. In the meantime, a program using a behavioral science approach to the problem of interpersonal relationships had been underway at the University of Chicago by a group of which Dr. James Grier Miller was the spokesman. The work of this group seemed appropriate to our setting at The University of Michigan. With this in mind, the idea was presented to the Administration of the University and then to some members of the Legislature in Lansing who were particularly interested in the problem of mental disease. The idea received strong support from the University and from members of the Legislature. Negotiations then were undertaken which led to the move of the group to Ann Arbor to become senior members of the staff of the Department of Psychiatry. Thus, a plan was developed setting up the Mental Health Research Institute with a nucleus of eminent scientists. Dr. Miller became Director of the group and Dr. Ralph Gerard became Director of the Laboratories. Included in the plan was Dr. Anatol Rapoport and others.

As the program for various research activities began to develop, it was obvious that more space for such studies would be required than had been originally anticipated. When this was made known to the Legislature, one member of the Legislature in particular, Senator Beadle, suggested that it would be worthwhile to request an appropriation to make adequate facilities available. With this encouragement, an approach was made to the Federal Government for aid by providing matching funds. About the same time, the McGregor Fund of Detroit contributed generously for special equipment. Our plan has come now to fruition. It seems apparent already that the research being undertaken will be worth many times the investment required to make such investigation possible. It is my belief that the wisdom of the Legislature in making available this means of doing real research has already begun to show evidence of considerable value and that in the future this return will be multiplied many times over. To quote from a recent article by de Takats, who expresses my sentiments so much better than I can:

> "Everybody agrees that the creation of a scientific atmosphere in a medical school or hospital, which can attract young investigators who are free to pursue their research interests without having to work on stated projects, is the best way to advance knowledge. In such a well-supported institution, the cross-fertilization of ideas, the 'fecundity of aggregation,' as Professor Walter B. Cannon called it, is the greatest way for one research worker to stimulate the other even if their fields of interest are miles apart. That is not integration, but 'juxtaposition of creative intellect,' as Means pointed out in his recent monograph, which every director of clinical research should read over and over again. The support of career investigators tries to maintain continuity of categorical research. The support of a research unit, a ward, a laboratory, is not popular today; it is the optimal way to spend money and see it well spent."

It is our studied opinion that the investment made by the State in this laboratory for research will make such a research unit more "popular" in a real sense of the word

Some months ago when the retirement of our former dean, Dr. A. C. Furstenberg, became imminent, a committee was appointed to recommend his successor. I had the good fortune to be a member of the committee. It was very soon realized what all of us had known before—that it would require an extraordinary person to continue to carry the responsibilities of the Medical School in the pattern established by Dr. Furstenberg. As a result of much investigation and a careful evaluation of potential candidates, the Committee recommended to the Administrative Officers of the University and the Regents of the University the appointment of Dr. William N. Hubbard, Jr. The wisdom of this recommendation is already apparent. He has undertaken his duties as Dean so actively and so successfully, and has fitted into the role so comfortably, that it would seem to be decidedly inappropriate to refer to him any longer as our "new" dean. It is with much pleasure, therefore, that I present to you Dr. William Hubbard, Jr., who will speak on the subject "Mental Health Research in the Medical School."

Dr. Hubbard: In the past twenty-five years there has been a spectacular increase in the research carried on in the medical schools of this country in both the fundamental biological and clinical aspects of problems of human health and disease. It is anomalous that during this period of remarkable increase and effectiveness of investigation the single most pressing health problem in the United States today—that is, the problem of disorders of behavior—has received relatively scant attention. It is interesting to speculate on some of the causes for this contradictory development.

Since the well-known report of Abraham Flexner was published in the year 1910, medical schools have seen their programs reoriented toward a university ideal of education as contrasted to the earlier emphasis on practical training. This has meant the introduction of curriculum changes which allow the medical student to work in an atmosphere more nearly that of the graduate school of the university. Collaterally, it has meant the encouragement and support of research activity by the faculty and the students as well. In order that its primary purposes of teaching and research can be accomplished, a medical school faculty has, of necessity, to render a considerable volume of direct clinical service to patients and that of the highest possible quality in order that optimum patient care may be demonstrated to future practitioners of medicine. Finally, where new or highly specialized techniques of patient care were developed in the university hospitals, it became their responsibility to make those new modalities as widely available as possible in the shortest period of time. This required patient service beyond that immediately needed for teaching and research and as well placed upon them the obligation of a considerable program of postgraduate medical education.

In order that these broad functions could all develop without undue competition one with the other, one finds the classical Greek concept of balance and symmetry being inherently meritorious appearing once again as a guiding principle of academic policy. As long as the subjects of inquiry were easily related to the existing departments and divisions of the medical school, the dynamic balance between development and symmetry was possible of achievement. Even here, however, this notion of balance was utilized from time to time to stultify and discourage major increases in research activities that were not accompanied by similar increases in teaching activity. Again, this has most commonly reflected a deficiency in the total amount of money available to support facilities and staff the ambitions of both the teaching and the research programs. The British notion of "fair shares" has been dominant.

In departments of psychiatry, the pressure of demand for direct clinical services is probably greater than in any other segment of the medical school. Quite typically, the number of patients for whom a department of psychiatry is responsible is far in excess of that number required for the discharge of its basic university responsibilities of teaching and research. As a result one finds these university departments in the medical schools engaged in large service programs

which have no necessary relationship to a university structure. The community's great manpower needs in psychiatry have resulted in rapid and appropriate increases in graduate and postgraduate training responsibilities for these same departments.

What seems to have happened is that the teaching responsibilities have been easily recognizable and traditional while the service obligations are apparent and vital to the conduct of the teaching effort. Less easily defined, newer in its demand for equal position in this triad of teaching, research, and service and competing for an already inadequate number of dollars, has been the area of research. When one recognizes that major improvements in health service depend immediately upon improved knowledge and that the creation of this new knowledge is the particular responsibility of a community of university scholars, then the logical basis for specific support of research is clear. It is only very recently, however, that the beginnings of practical achievement of this kind of support has been evident.

The first of the impediments that I visualize to the development of research in mental health in the medical school has been the overemphasis on the merits of balance in the medical school programs. Second, is the nature of the basic disciplines brought to bear upon the problems of mental health.

The theoretical structure of analytic psychiatry has not lent itself easily to the kind of quantitative critique that developed first in the physical sciences and then was adopted in experimental biology and medicine. The difficulty of describing the phenomena of behavior in terms understandable and acceptable to other clinical and basic scientists has made it difficult for the problems of mental health to be brought into the main stream of scholarly activity in the university. This is, of course, not unlike the relationship of clinical medicine to the other natural sciences at the turn of this century. It now seems probable that the most fruitful approaches to research in mental health will demand the representation of many disciplines not traditionally a part of a medical school faculty. On the other hand, these newer contributions will be of very doubtful effectiveness if they are not in turn related closely to the broad spectrum of interest in human health and disease represented in the traditional medical school structure. In this, mental health is not alone—and physics, mathematics, engineering, anthropology, law, and theology are all active participants in other areas of the medical schools' efforts. As tools for quantitative investigation of problems of mental health become available, the workers in this field draw closer to other investigators in the biological and physical sciences. As the ideal of the university, as representing a community of scholars, receives increased emphasis, it becomes possible to develop broad areas of investigation which do not conform to traditional departmental distinctions.

The third impediment that I would suggest, after the problem of balance and that of quantitative and broadly representative tools of investigation, is the matter of interrelationships within the university family.

Geographic proximity surely is an important characteristic of unity in the university family. Many schools of medicine in the last fifteen years have followed the precedent set by The University of Michigan many years before and established their schools in relationship to the University rather than (as traditionally) in contact with a large pre-existing hospital. The proper academic bond, however, between the institutional parent and her offspring, the medical school, is dependent upon many other factors—some of them much less tangible than geography. The pure academician, despite frequent outward denials, may carry a subconscious feeling of separation from the professional school. To him, a school of medicine may appear to be primarily a training institution—a trade school with technical emphasis—rather than an integral segment of a university family devoted to higher learning. In the past years, and even today, there may be some valid reason for this judgment. Medical education does, in fact, find itself confronted with the absolute necessity of

training physicians who are essential to the health of the community. Furthermore, those physicians who work in the University Hospital carry all the traditional responsibilities of physicians for their patients. As a result, to the pure academician it is not clear whether the allegiance of such a faculty member is to a hospital or to a university. No one denies the importance of the basic teaching program or the hospital responsibilities, yet they have amounted to a stigma when the academic program of the medical school is compared to that of the graduate school. As the quality of scholarly activity increases in the medical school and as research becomes more closely interwoven with teaching and service, it is my conviction that this problem of interrelationships within the University community itself will be largely resolved. Mental health research in the medical school, therefore, is a particular and very important instance of a more general problem of the proper development of medical research. The interdependence of teaching, research, and patient service, if excellence is to be achieved, is well represented by the organic relationship of this Institute for Research in Mental Health with the Department of Psychiatry. The need for balanced emphasis in the institution has been recognized through the provision of independent and noncompetitive support of this activity in research.

As one looks around at the problems of the aging population, disorders of cellular development and cancer, so-called nuclear medicine, human genetics, and the newer fields of aviation and space medicine, this example in The Mental Health Research Institute of special support for interdisciplinary effort in selected fields, while maintaining close and active co-operation with the basic programs of the University and Medical School, achieves the status of an important precedent. It is not an Institute conceived as being separate from other activities of the University faculty. It does not cut itself off from the continual flow of students that constitute the next generation of investigators in this and other fields. It does not withdraw from the responsibility of introducing to future practitioners of medicine the concepts which brought this interdisciplinary institute into being.

On the other hand, it is possible in this and in other such institutes to gain some release from the restrictive intimacy of responsibility for clinical phenomena. In such a position of freedom it is possible to move from the study of simple cause-effect relationships that have dominated thinking in the past. The conditioned reflex in its simplest form and such simple cause-effect concepts as the relationship of thyroid function to cretinism can be used as a stepping stone. Already, the actuarial approach—or the use of probability statements—has begun to introduce the statistical notion of probabilities in randomized effects rather than restriction of concept to simple first order reactions. Just beginning is our understanding of the effect of complex associations that are not randomized but are organized instead. The dynamic relationship between the individual and these units of organized complexity represents an exciting and productive field of investigation for biologist, clinician, mathematician, and engineer.

There is no doubt that mistakes will be made in the direction of this newer kind of research. There is, however, one sure way to avoid mistakes and that is to do nothing. Or at least to do nothing new. This, of course, for the scholar is the greatest mistake of all.

The role, then, of mental health research in the medical school is first to bring the scholars of the University together in a setting which will encourage the development of that new knowledge which is the only source of a solution for the most pressing problem of health in the United States. Beyond this, however, effective mental health research will broaden classical approaches to human biology by emphasis of mathematics, engineering, and the behavioral sciences of sociology, psychology, and anthropology.

Furthermore, active research in this field will broaden the orientation of classical clinical psychiatry with the introduction of the concepts of electrophysiological, bio-

chemical, pharmacological, histochemical, genetic, and refined morphological studies.

All of this, then, will be done in an atmosphere open to competent and interested students and investigators who can share together the spirit of adventure that is inherent in a search for understanding.

Dr. Waggoner: Thank you very much, Dr. Hubbard, for a very interesting and matter-of-fact discussion of your subject. You have pointed out most effectively that it is difficult to present in a concrete fashion research studies being done in a field so esoteric as ours.

Real scholars are most difficult to come by and rare indeed are scholars who have accomplished so much or demonstrated their productivity so well as our next speaker. Dr. Ralph W. Gerard, of the Mental Health Research Institute, has so many accomplishments and honors in his record as evidence of his scientific eminence that I shall not even attempt to enumerate them but will simply ask him to present to you, in a necessarily brief fashion, the program of the Institute.

Dr. Gerard: My assignment, to present the program of our Institute in less than half an hour, makes me think of the closing line of Christopher Morley's "To a Post Office Inkwell"; the last verse is:

> "Your pool of ink, your scratchy pen,
> Have moved the lives of unborn men,
> And watched young lovers, breathing hard,
> Put heaven on a postal card."

But I shall try!

Lord Adrain, one of the world's great scientists, is also master of Trinity College, Cambridge. A few years ago, when three Nobel prizes were awarded to Trinity Fellows in the same year, he commented, at a small celebration honoring them, somewhat as follows: "Of 36 Nobel Laureates in the entire United Kingdom, 13 have been Fellows of Trinity College. It would be unbecoming for us to be prideful over this record—but in any other institution this might be forgiven."

Trinity had then been in existence just over four centuries. We have been in existence just over four years and have as yet not much to be prideful about; but, as the gardener at Buckingham Palace said to the tourist who asked how he obtained such a fine lawn, "You just roll it and you cut it and in a few centuries it gets that way."

We are, in fact, still amorphous and a bit squashy—like an embryo; and, also like an embryo, we need much and increasing nourishment. We have been fortunate so far, with good support: from State and University; from United States agencies, such as Army, Air Force, especially Public Health, and others; from the McGregor and other private funds. We have a capable and enthusiastic young staff. And we have a good program.

Our *program*, as it should be, is in line with the charter given us by the legislature. This specifies four charges, on which I now report to you.

The first is to consult on state mental health problems. Some of us are on various Governor's committees dealing with mental health; some have participated in specific conferences of the Department of Mental Health; several have worked closely with the state mental hospitals. Not only on the state but on the national scene has consultation been extensive; Institute members have served on advisory committees and panels and boards for the National Institutes of Mental Health of the Public Health Service, the Veterans Administration, the National Science Foundation, the National Research Council, and other agencies. Our members spearheaded a special committee on Behavioral Science requested by Vice-President Nixon and are serving on the only official national committee working in this area.

The second charge is to teach students (including residents) and to encourage research. Four courses have been instituted and given to residents in psychiatry; courses or lectures have been given in a bewildering array of departments, including: architecture, art, history, psychology, economics, engineering, conservation, zoology, biochemistry, physiology, pharmacology, dietetics, nursing, and so on. Over 130 students have worked with us, doing research of their own or acting as assistants in an ongoing research program. These included half a

dozen residents, over a dozen graduate students, half a dozen medical students, two dozen undergraduates, and over seven dozen part-time technical helpers from various levels. Institute members have consulted and co-operated with many other members of the Department of Psychiatry in initiating or pursuing their own research programs.

A third charge is to hold conferences. The Institute has organized specific conferences on History and Social Change, Management Simulation, and Social and Behavioral Science; and Systems Theory Workshops have continued at the rate of three or four a year from our inception and have involved a couple of dozen national leaders in various relevant fields. Many University lectures have been arranged; the Institute has a regular series of weekly seminars on topics within its purview; and it has offered four special teaching seminars on subjects ranging from neurophysiology to computer programming. The quarterly journal *Behavioral Science* was founded by the Institute and now reaches well over 3,000 subscribers.

But, of course, our main charge is the last one—to do research ourselves, and especially in the general field of mental health. Before I can comment meaningfully on this large portion of our program, a few words may be helpful on our picture of mental health.

Mental health includes far more than the individual's psychic welfare and the elimination or treatment of psychoses and neuroses. It involves, as well, helpful interactions between men in groups, small and large, and in their total societies. It depends on the action of the organs and cells and molecules in individuals as well as on the social interactions between them. It constitutes, in fact, a very large sector of the whole domain of behavioral science.

Behavioral science has come to have various meanings to different people. Our picture has behavioral science including a larger or smaller chunk of many of the existing disciplines, especially in biological and social science. If we think of the existing departmental disciplines—physiology, psychology, sociology—as vertical columns, then behavioral science is a group of horizontal task forces, each composed of recruits from a selection of departmental areas and each directed to the solution of some particular problem. Indeed, this is a repetition of the earlier history from which the present disciplines developed.

Man was early faced with problems of land apportionment, and developed geometry; with problems of food raising, and developed botany and agriculture; with problems of health, and developed anatomy and physiology and pathology. These became sciences in their own right, as an adequate body of knowledge and interpretation coalesced; and from them, in turn, came new interdisciplinary solutions to problems. Thus, primitive medicine was an art; from it came the life sciences; and from their collective application came clinical science. Modern medicine is thus, itself, a behavioral science.

As science solved the environmental problems, first in the physical region and then in the biotic one, man was able to increase in numbers and in concentration in urban communities, until today the overriding problems that face mankind are those of his social environment—of man's interactions with other men. Hence the overwhelming importance, now, of advance in the field of behavioral science.

Of the horizontal task forces that constitute collectively behavioral science, mental health is involved centrally in some of the most important, and heavily in a great many others. Psychoses and neuroses, which involve as much as a third of our population; addiction of all sorts, including alcoholism, today our major health problem; delinquency, the most rapidly rising threat to internal social stability; and the not unrelated accelerating divorce rate, are all directly problems of mental health. But mental health is perhaps the dominant, certainly an important, element in such areas as aging, medicine, public health; in problems of traffic, stress, social acculturation; in the improvement of communication,

education, architecture, law; and so on and on through most of the institutions and values of human culture. It is the core of war and peace, of conflict and conflict resolution.

Behavioral science implies a faith, necessarily held by all behavioral scientists, that human behavior, even its irrational components, can be understood; and, further, that men with different trainings can work together effectively toward such an understanding. This is the very core of the Institute's mission and our research is set in this framework.

Our various *research* projects are presented, probably in sufficient detail, in the Third Annual Report of the Institute, available to all. Ranging from biochemical studies of nerve cells under the microscope to a consideration of the impact of one civilization on another, our projects may seem widely scattered and hodgepodge. That they are not so kaleidoscopic as may appear, a little consideration of the organization of knowledge will show.

The various entities or *systems* that constitute our universe of interest clearly fall at different levels, forming a hierarchy of progressive inclusion and increasing complexity. Let us arrange them as a vertical series, from molecules through cell, organ, individual, small and large groups, to society—as shown in Chart 1. Next it turns out that the entities or systems at each level have alike three major attributes: their architecture or structure or morphology or organization; their functioning or behaving or responding; and their history or evolution or development or learning. Being addicted to alliteration, I have christened these attributes: "being," those structural properties which remain constant through time; "behaving," involving reversible changes; and "becoming," standing for the irreversible developmental aspects. These are shown as a horizontal row in Chart 2. Of course, no scientist worth his salt can resist making a table when he has two axes like this facing him, and Chart 3 shows the resulting squares or boxes, vertical columns and horizontal rows, that emerge when such a table is drawn.

There is not time today to indicate the great value of such a formulation; it is a

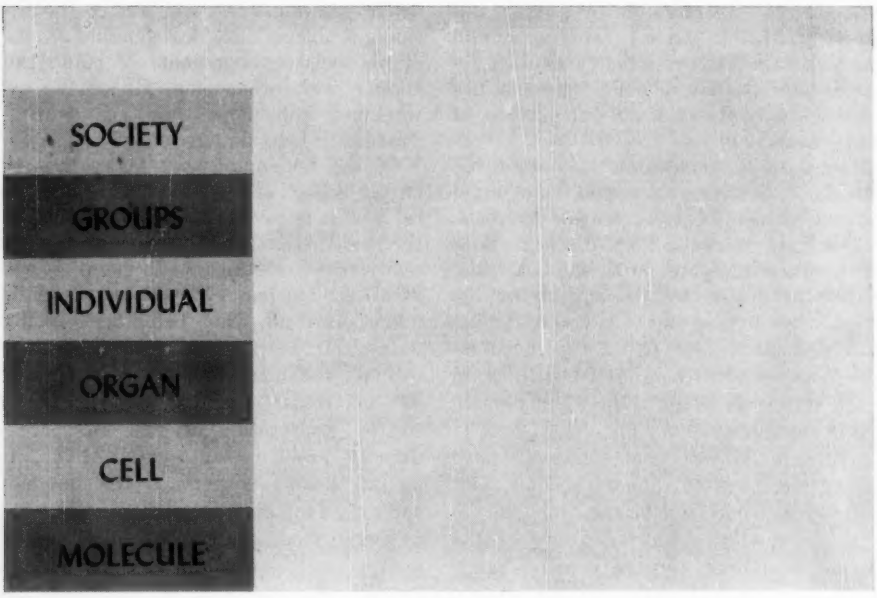


CHART 1

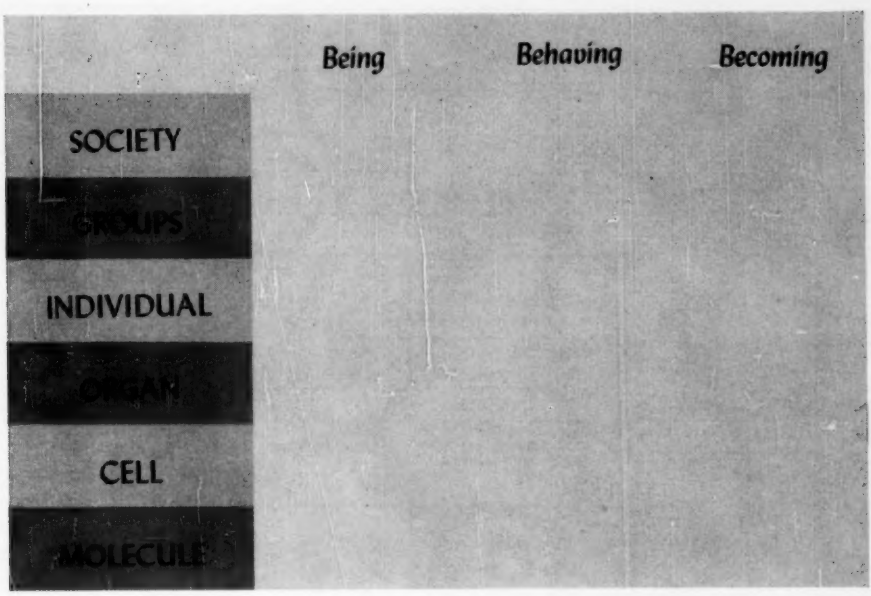


CHART 2

	Being	Behaving	Becoming
SOCIETY			
GROUPS			
INDIVIDUAL			
ORGAN			
CELL			
MOLECULE			

CHART 3

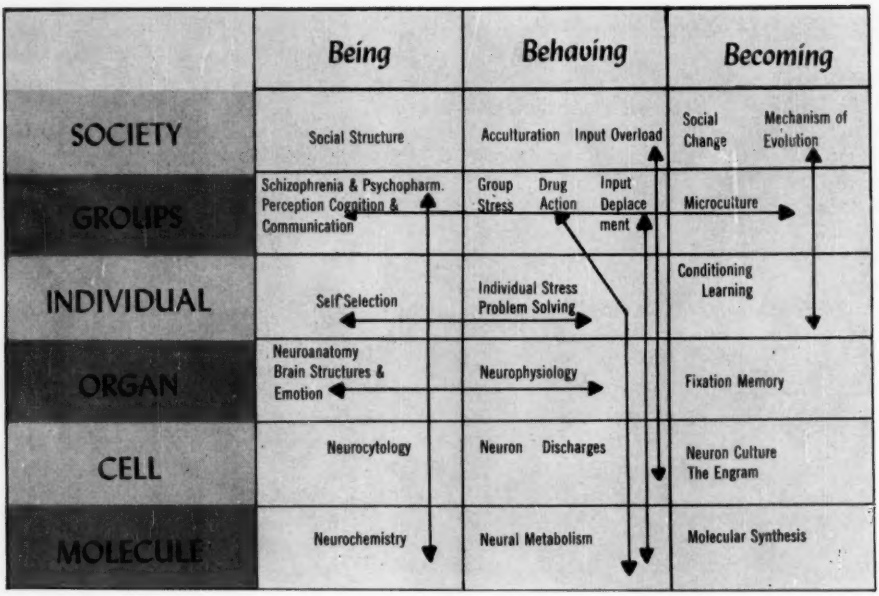


CHART 4

powerful intellectual map for understanding the historical development of science, the spiral relations of experience and structure and function, and the differences between units grouped by origin, like species; and by function, like communities. It is a stimulus for identifying the correct basic units for scientific study, especially in the "behaving" and "becoming" columns, and for the discovery of common principles that are valid for all levels of systems, by examining the vertical columns or their subcolumns. And it exhibits the complex relations of a system at any given level to its superordinate and subordinate systems—the problem of individual freedom verses group co-operation is a characteristic one.

I can only gloss over Chart 4 on which is indicated the various recent or current investigations at the Institute. Some of the major programs will be detailed by speakers in the afternoon; the largest single one, a co-operative research between the University of Michigan and the Ypsilanti State Hospital, on schizophrenia and psychopharmacology, must be passed by; but I can say now that on this project some real progress is being made in revealing the architecture and meaningful characteristics of this major mental illness.

Time presses and I must shortly close. We are only four and a half years old but, even so, there is a little to show for our existence. The educational activities have already been mentioned. Something over 150 articles have been published, over half in this past year, and more then a dozen books have been published or are in press and several more are in preparation. Surely one or two dozen of these are of some significance.

Three original homeless space wanderers, with a secretary, have, after eight moves, come to rest in a fine 50,000 square foot building with, as the airplane manifest quaintly puts it, over 120 S.O.B.s—souls on board. Our three floors were likened, by a visiting British physiologist, to the three great English Universities. On our first floor are conference and lecture rooms,

interview and group research facilities, administrative offices; upstairs are our library and computer rooms, lounge and lunch rooms, and our studies; and below, on the ground floor, are the "wet" laboratories, the microscopes and oscilloscopes and centrifuges and animal quarters. Our visitor suggested that the first floor be named for Oxford, where they talk; the second floor for Cambridge, where they think; and the ground floor for London, where they work!

We are grateful for much and generous support from many, notably from Ray Waggoner who has helped and encouraged us unreservedly. Support has come widely from many—largely, I hope, present at this dedication—who believe in the Institute and its potential; we have even had support from some, also present I hope, who have entertained serious concern as to our becoming delinquents.

We have moved from a handful of men who believed deeply in the need for and possibility of a better and a more rational handling of the mental health of mankind to an enthusiastic group of men and women now bending their efforts toward achieving this.

Science has doubled every fifteen years over the past three centuries. Time rushes ever faster. Our sights are on Trinity College, on the Rockefeller Institute, and on other great institutions. We hope that, not in 400 years, but in forty—we dare dream, in moments of recklessness, that even in four—we will have achieved something worthy of this goal.

Ladies and gentlemen, this is *our* dedication.

Dr. Waggoner: Thank you very much, Dr. Gerard. Your delightful, brilliant, and clear description of the work of the Institute gives us a bird's eye view of the program of this new activity. It has been most heart warming to have such a representative group meet with us for these exercises, and I should like to express our very real appreciation to you for your attendance.

* * * * *

The afternoon session was presided over by Dr. Donald G. Marquis of the Massachusetts Institute of Technology, formerly a member of the MHRI Advisory Committee who participated in many of its staff activities in the early, formative years. This last session of the day was devoted to presentation of three scientific papers, representative of the interests of the Institute. Much of the material in these papers has been published elsewhere. Dr. Jacob Marschak of Yale University spoke on "The Economics of Leadership"; he plans to publish this material at a later date. Dr. Anatol Rapoport of the Institute talked of his research on "An Information Theoretical Approach to Individual and Group Learning" (see Rapoport, A. A derivation of a rote learning curve from the total uncertainty of a task. *Bull. Math. Biophy.*, 1960, 33, 85–97), and Dr. James G. Miller, Director of the Institute, discussed his work on "Information Input Overload" (see Miller, J. G. Information input overload and psychopathology. *Amer. J. Psychiat.*, 1960, 116, 695–704).

Economics is the science of phenomena due to one love and one aversion—gain and labor.

JOHN DEWEY, *Human Nature and Conduct*

CRITIQUE AND COMMENT

Any organization is happy with—and likely to promote—a man who is wise enough to make correct decisions for its successful operation. But in a large bureaucratic organization, may not statistical probability alone result in there being at least one man who makes all the correct decisions but who is no more truly competent than his less successful peers? How do we distinguish the wise from the pseudowise? These intriguing questions and some of their ramifications for political and organizational behavior are here discussed.

A NOTE ON THE APPEARANCE OF WISDOM IN LARGE BUREAUCRATIC ORGANIZATIONS

by Karl W. Deutsch and William G. Madow[1]

Yale University and Stanford Research Institute

MODERN life is characterized by the existence of many bureaucratic organizations, such as government agencies, the military services, political parties, educational institutions, and business corporations. In such organizations, officials sometimes are called upon to make decisions which are considered important, and which are recognized later as having been clearly right or clearly wrong. It is often customary to single out for praise or promotion those officials whose decisions in important matters have frequently or always turned out to be right. The presence of such tested and proven wise men on its decision-making staff may then strengthen confidence in the wisdom of the organization that employs them, since it may be thought likely that these wise men and their organization will continue to make correct decisions in the future.

Such expectations sometimes may be justified and sometimes not. The correct answers may have been arrived at with the help of some manifest or hidden resources or procedures which in fact are likely to continue to yield correct results when new decisions must be made; or else they may be the product of a blind statistical process which may leave the apparent wise men and their employers helpless at the next juncture.

It seems desirable, therefore, to try to separate these two phenomena: (1) the instances of presumably genuine wisdom, which we may define for the purposes of this argument as the likelihood of an individual or small group to make correct decisions, for reasons inherent in the actor; and (2) the statistical appearance of pseudowisdom, defined here as the accidental making of correct decisions in the past (e.g., for reasons external to the actor), and unlikely to lead him to further correct decisions any more often than his previously less successful peers. If we can say something about the probable extent of the appearance of such pseudowisdom in large organizations under various conditions, it should be easier to say whether the decision makers in the organization have performed about equal to statistical expectation, or whether they have done significantly worse or better.

For purposes of exposition of the argument that follows we make the following as-

[1] The authors are indebted to the Center for Advanced Study in the Behavioral Sciences where they spent the year 1956–57 as Fellows and where most of the work on this paper was done; and to Dean W. Allen Wallis and Mr. Walter Pitts for helpful suggestions. The responsibility for the present paper remains, of course, their own.

sumptions: (1) there is a number of clearly identifiable bureaucrats in the organization who are charged with the making of decisions and whose decision-making records can be compared. Let the number of such bureaucrats be m; (2) there is a number n of important decisions to be made which can be clearly distinguished from trivial decisions; (3) it is further assumed that these important decisions are independent of each other; (4) it is assumed that the organization has a clear criterion for distinguishing "correct" from "incorrect" decisions within a reasonable time after the decisions have been made, so that the past decision-making performance of each bureaucrat can be clearly established.

Assuming that a reputation for wisdom will be established by a bureaucrat who makes correct decisions in a limited number of important cases, it should follow that in a large bureaucratic organization a number of bureaucrats will establish such reputations, even if their probability of finding the correct answer in each case were no better than pure chance, such as .5 or the fall of a true coin. If we call the probability p of a bureaucrat making a correct decision his "competence," and if we assume for the sake of argument that this level of competence is the same for all bureaucrats in the organization, we may ask the question: How many apparently wise men will be produced by pure chance for a given number of bureaucrats and of important decisions, and for a given level of bureaucratic competence?

A SIMPLE EXERCISE IN PROBABILITY

Let p be the average probability of making the "correct" decision in a single instance. Let n be the number of independent decisions to be made, and suppose that correctness in all n decisions is required of an individual bureaucrat in order to give him a reputation for great competence. (Thus, if $n = 8$, there are eight independent decisions to be made and a record of eight correct independent decisions would be considered, at face value, to be evidence of high ability.) Then let there be m bureaucrats on the same initial level and suppose we want to determine m such that one bureaucrat is expected to emerge after n decisions with a perfect record of correct independent decisions.

Example 1:

Let $p = 1/2$
$n = 8$
m = number of bureaucrats at start of testing period
y = expected number of bureaucrats with perfect records at its end.

Taking $y = 1$, it follows that $m = 2^8 = 256$, or generally, $m = y(1/p)^n$. (Other conditions for determining m are discussed below.)

Thus, among 256 bureaucrats each having to make eight major decisions, there should be on the average one with a perfect record of correct decisions, even if his actual probability of deciding correctly was only 50:50 in each case, and even if nobody learned anything from any previous decisions.

Furthermore, the probability that at least one of the 256 bureaucrats would make all eight decisions correctly on this random basis is about 2/3, and even the probability that four or more of the 256 would make all eight decisions correctly would be about .05.

If the level of competence is less than .5, or if in each decision a choice between more than two alternatives must be made, a larger number of bureaucrats, m, will be required to produce the expected appearance of one spurious wise man.

Example 2:

Let $p = 1/3$
$n = 4$ $\qquad m = 3^4 = 81$
$y = 1$

In actual situations, bureaucrats usually are more competent than tossed coins. If so, p will be larger than 1/2 and the ratio of m to p will be much less. (Also, there usually is some learning; but the learning effects and the probability effects should be disentangled.) The case of somewhat more competent bureaucrats, with a 2:1 probability of making correct decisions, is illustrated in Examples 3 and 4:

Example 3:

$$\text{Let } p = 2/3,\ n = 4,\ y = 1 \qquad m = \left(\frac{3}{2}\right)^4 = \frac{81}{16} = 5.06$$

Example 4:

$$p = 2/3,\ n = 8,\ y = 2 \qquad m = 2\left(\frac{3^8}{2^8}\right) = \frac{6561}{2^7} = \frac{6561}{128} = 51.26$$

Thus, even a half-dozen bureaucrats may produce one "wise man," if only four decisions are considered relevant; and if we are looking for two bureaucrats with a perfect record on eight major decisions, we may expect to find them in the majority of organizations large enough to employ at the appropriate bureaucratic level more than 52 decision makers of average ability.

Following a similar line of reasoning for the field of private business enterprise, Professor Harold Guetzkow has suggested that out of a large number of new business firms, there should emerge an appreciable number of survivors y with a perfect record of correct major decisions per year or even per decade, so long as the number of "major" decisions, that is, n, remains small relative to m, the number of new firms.

This situation may occur not infrequently in real life, when the standards for the routine conduct of small businesses have become fairly uniformly accepted, and the success or failure of competitors may depend largely on a very few major decisions such as the choice of location, the choice of lines of goods to be carried, and the decision whether or not to seek substantial credit for expansion. If half a dozen major decisions of this kind should be sufficient to decide about success or failure among hundreds or thousands of competing small businessmen, statistical chance might contribute a substantial contingent to the ranks of those successful men who may be tempted to congratulate themselves upon their prudence and foresight in making the right decisions for business survival.[2]

Our model could be further refined by introducing z as a "coefficient of wisdom," that is, the smallest acceptable ratio of correct to total relevant decisions. In the preceding dicsussion, z was held equal to one; but if a "batting average" of 0.8 or 0.4 should be acceptable, the number of bureaucrats or businessmen with acceptable records should increase in accordance with some formula connecting z with the other variables.

At the present stage it should be more useful, however, to make explicit the effects on the size of m of taking account of statistical variability in the case of the simple models discussed above.

A FEW MATHEMATICAL NOTES

Let us recall and elaborate our simplifying assumptions. We shall suppose that every decision of each decision maker is independently made of all others that he makes, and that each decision maker is independent of all other decision makers. Furthermore, we shall suppose that there is a fairly uniform probability p of being correct in a single decision, which holds for all decisions and all decision makers. Clearly, this assumes an

[2] Dean W. Allen Wallis has drawn our attention to a related example from the practice of stock market fraud. According to Mr. Wallis, the practice consists in sending letters about some highly speculative stock to, say, 4000 prospective investors, advising half of them that the stock is certain to rise, while telling the other half that it is certain to go down. After some weeks, to those 2000 investors to whom "correct" advice was given as proved by subsequent events, a new letter is sent about another speculative stock telling again one half of the addressees that the stock will rise, and asserting the opposite to the rest. This practice is then repeated for the thousand who have now received two "correct" predictions in a row, and to the 500 who received three correct predictions and to the 125 who received four correct predictions. At this stage the authority of the forecasting service may be so thoroughly established with this last group that they may be willing to part with substantial amounts of money to buy some worthless stock which is now offered them at the fifth stage, by the same source whose authority has been so impressively demonstrated to them.

excessive amount of uniformity and mediocrity. Yet these conditions have been chosen to highlight the process by which appearances of singularly high performance are produced. The same tendency will occur when the probabilities of being right vary among decision makers and decisions.

The probability that a specified decision maker is correct in j of n decisions is then

$$P(j, n) = C_j^n p^{j} q^{n-j};$$

and the probability that all n decisions are correct is

$$P(n, n) = p^n,$$

which we shall denote by P.

If there are m decision makers, then the probability R that at least one of them makes all decisions correctly is

$$R = 1 - (1 - P)^m = 1 - (1 - p^n)^m$$

since $1 - P$ is the probability that any specific decision maker makes at least one error.

Now suppose we want to fix n at eight and to determine m so that R has some specified value, say, .95.

Then, the number of decision makers m_n such that the probability is R that at least one of the decision makers makes all n decisions correctly, is

$$m_n = \frac{\log (1 - R)}{\log (1 - p^n)}$$

and the necessary values of m_n for specified values of p are given below.

Similarly the probability that at least $n - 1$ of the n decisions are correctly made by a single decision maker is

$$np^{n-1}q + p^n = p^{n-1}(p + nq)$$

and hence the probability that at least one decision maker is correct in at least $n - 1$ decisions is

$$C = 1 - [1 - p^{n-1}(p + nq)]^m$$

and, the number of decision makers m_{n-1} such that the probability is C that at least one of the decision makers makes at least $n - 1$ decisions correctly is

$$m_{n-1} = \frac{\log (1 - C)}{\log [1 - p^{n-1}(p + nq)]}.$$

Many extensions of these results may be obtained.

Furthermore among m decision makers the expected number y of decision makers who make all n decisions correctly is

$$mp^n$$

and the variance of the number of decision makers who make all n decisions correctly is

$$mp^n(1 - p^n).$$

It is worth remarking that the Poisson approximation to the binomial will be good if m is fairly large, since even $(1/2)^8$ is fairly small, i.e., 1/256. Also, if one wishes m to be large enough so that the expected number of decision makers who make all decisions correctly is an integer K, e.g., 1, then one determines m by the equation

$$m = \frac{K}{p^n}$$

and hence m will often be large enough for the Poisson approximation to be used, unless n is very small.

Let us finally consider a related problem. Suppose that a decision maker has a probability P of succeeding (e.g., success may mean that he makes eight decisions correctly), and that we test successive decision makers until K are successful. (This might be done, for example, to choose a team of K "successful" decision makers.)

Then, the probability that the Mth decision maker is the Kth successful decision maker is

$$C_{K-1}^{M-1} P^K Q^{M-K}, \quad M = K, K + 1, \cdots$$

Also the expected number of decision makers that must be considered until K are successful is

$$\frac{K}{P}$$

and the variance of the number of decision makers that must be considered until K are

successful is

$$\frac{KQ}{P^2}$$

where $Q = 1 - P$.

SOME IMPLICATIONS FOR POLITICAL AND ORGANIZATIONAL BEHAVIOR

Some effects of size of organization

The general shape of these probability functions shows that selection and promotion of high-level personnel in large organizations may be seriously affected by the contaminating effect of the appearance of spurious "wise men," if the number of competing bureaucratic decision makers is large, while the number of important decisions is small. Thus, every major country has several hundred generals in its armed forces, yet each of them might have to make no more than half a dozen really crucial decisions in the course of his career. If so, several "military geniuses" might be produced by chance, and these men might be put into important positions without being in any way more competent than those other men over whom they are preferred.

Some effects of the steepness of personnel selection

This contamination effect itself depends only on the number of crucial decisions and on the number and competence of bureaucrats at the lower level of promotion. Thus, in our first example, 52 bureaucrats facing eight crucial decisions with indifferent or 50:50 competence would produce an expected number of two spurious "wise men," regardless of the number of higher-level positions to be filled. The possible impact of this spurious effect on the organization may vary, however, with the proportion of personnel selected for promotion. If 10 of the 50 were to be promoted, the spurious "wise men" could be expected to fill no more than two of these 10 positions, or one-fifth of the total at the higher level. The contamination effect at this higher level would thus remain relatively limited.

If, on the other hand, only one or two of the original 50 were to be promoted, then the situation might become more hazardous for the organization involved: sometimes the spurious "wise men" might happen to fill one-half or even all of the positions on the next higher level. The situation would be similar to the taking of small samples from a large universe: some of them would represent the results of variance and not the average composition of the system as a whole. In this way a few small samples may be unrepresentative of the total; and gamblers with limited funds risk being wiped out by fluctuations of the game even if the long-run odds were in their favor. The same may hold true for the problem of high-level personnel selection in large organizations. Analogous to the gambler's ruin, the possible impact of the contamination effect at higher levels will thus increase with the steepness of selection, while the long-run probability of spurious "wise men" being picked instead of real ones will continue to vary with the relative proportions of the former to the latter.

A bureaucratic organization thus should become more vulnerable to the contamination effect on its higher levels in proportion to the steepness of its organizational pyramid. The smaller the proportion—let us call it s—of higher bureaucrats at each level to those immediately below them, the greater—other things being equal—would be the effect of variations from the average, and thus the greater could be in some instances the actual proportion of spurious "wise men" in the higher echelons. This risk becomes more significant as s, the proportion of bureaucrats to be promoted, approaches or falls below y/m, the proportion of bureaucrats likely to appear wise by chance.

In such cases, what could a large organization do? At best, it could insist on evaluating a large number of candidates for each high-level position, even where a few apparently well-qualified ones are already in appearance. This would require a greater expenditure of time and resources on the evaluation of candidates from outside the "inner circle," and it could reduce the expected contamination effect to its original proportion on the lower bureaucratic level,

but it could not make it any smaller than that.

Some effects of the difference between trivial and crucial decisions

Are there more general ways in which the statistical contamination effect in large organizations could be reduced?

One remedy would consist in increasing the number of crucial decisions. One way of doing this is to wait longer, until more crucial decision situations had occurred, but this may often not be practical. Another solution would be to treat trivial decisions as if they were samples of important ones. Thus, big business organizations have at times been satirized as being inclined to take the good judgment which a junior executive has shown in the choice of his neckties as an indication of his presumable judgment in the making of major business decisions. The satirical description of this practice may be much exaggerated, but our previous remarks may have suggested that the practice itself may not be wholly without reason. Nevertheless, trivial decisions cannot be always presumed to be samples of critical ones: Ulysses Grant seems to have made an undistinguished record in regard to the routine that marked the career of an ordinary Army officer, but showed remarkable skill in making some of the crucial decisions that led to the winning of the Civil War.

From this we may derive a general proposition: *the greater the difference between routine decisions and crucial decisions in an organization, the more vulnerable this organization is likely to be to the contamination effect* that we have been discussing. The relative distribution of crucial versus routine decisions in any organization is to a large extent determined by the kind of work which it does. Thus, doctors in a hospital, or field commanders in wartime, may be called upon more often to make crucial decisions than are file clerks in a bank or post office. This does not mean, however, that the Army or a large hospital will always have a better chance of distinguishing the reality of talent from its spurious appearance. What counts will be rather this: how large will be the number of decisions that will be relevant for the reputation or promotion of a bank manager or post office superintendent, as against the number of decisions relevant for the career of a chief surgeon or a colonel? The larger the number of relevant decisions which the candidate has to make in establishing his record, and the more similar the nature of these decisions to the crucial decisions which he will have to make in behalf of the organization in the future, the safer the personnel policies of the organization will be from the effects of statistical contamination. Up to a point, organizations can deliberately increase the number of crucial or at least non-trivial decisions with which they are confronting a candidate for promotion before they decide about his career.

On the other hand, organizations, wittingly or unwittingly, can do the exact opposite. They can reduce the effective number of decisions that are crucial for a man's career, and thus they may make their own task of selection of able personnel more difficult and the promotion of incompetents more likely.

Some effects of ideological tests and political purges

A powerful way of increasing these undesirable effects consists in the introduction of political, religious, or ideological tests, and of political purges. Such tests, if taken seriously, usually will tend to override in importance any criteria of technical performance. The number of such independent political tests will be few, corresponding to major changes in the governing personnel or in the ruling ideology or party-line—changes which in the nature of things must be relatively infrequent, occurring on the average less than once a year. During the ten or fifteen years decisive for their careers, bureaucrats under a totalitarian system of government are not likely to go through more than perhaps six or eight major purges or changes in the party line. Yet the decision of the bureaucrat as to which side to choose at each such juncture, and how to forecast correctly the winning side, may be of much greater importance for his fortunes and his

reputation within the totalitarian context than any decisions on merely technical matters.

Out of the several hundreds and thousands of middle-level bureaucrats, a totalitarian system may thus tend to promote a number of persons who are genuine artists in well-timed conformity, and a number of others who have demonstrated genuine high competence in their technical fields, and even a small number of persons who rank high in both opportunism and competence. But on the showing of our argument, they will also tend to promote an appreciable number of persons of indifferent competence and loyalty to the currently ruling party line or faction. These will be the persons who chose sides with indifferent skill in each of the crucial political purges but survived them in accordance with the laws of probability.

Totalitarian or quasi-totalitarian personnel practices may thus be less selective than they seem, and the introduction of ideological tests or purges in the large bureaucratic organizations of free countries might make their personnel policies rather more erratic and unreliable than their proponents would intend.

Some possible effects of "emergency thinking" on public confidence in large organizations

More broadly speaking, whenever personnel selection is carried out against a background of "emergency thinking," it is part of the concept of an emergency that decisions made to meet it are considered much more important than the routine decisions that are being made at other times. If the workings of a government or of a large private organization should come to be characterized by a succession of spectacular emergencies, and if their bureaucrats should come to be judged mainly in terms of their performance in meeting them, the effects might be somewhat similar to those of a succession of major political purges. In both cases, a small sequence of spectacular decisions will become decisive for personnel selection and promotion while most other decisions made by the same bureaucrats would be relegated to comparative insignificance.

A related consideration should hold true for the reputation of large bureaucratic organizations with the general public. The larger the number of decision-making bureaucrats in such an organization and the fewer the number of decisions to which the public pays attention, the easier it should be for the organization or agency to impress the public with the appearance of time-tested wisdom among its high-ranking personnel.

(Manuscript received April 10, 1960.)

> M. de Roannez said: 'Reasons come to me afterwards, but at first a thing pleases or shocks me without my knowing the reason, and yet it shocks me for that reason which I only discover afterwards.' But I believe, not that it shocked him for the reasons which were found afterwards, but that these reasons were only found because it shocks him.
>
> PASCAL, *Pensees*

If in a large sample of written or spoken language the words are ranked according to the frequency with which they appear and the frequency is plotted against the rank, the resulting curve can be fitted by the equation $P = P_o(\bar{R} + B_o)^{-\alpha}$, where P is the frequency with which a group of words occur, $\bar{R}$ is their average rank, and P_o, B_o, and α are parameters. The equation has been rigorously derived by B. Mandelbrot on the basis of some theoretical considerations. In this derivation, the parameters P_o, B_o, and α receive a physical interpretation, which converts them with psycholinguistic concepts. Here the values of all three parameters are compared in several language samples, ranging from James Joyce's *Ulysses* through Standard English, to the language of a diagnosed schizophrenic with paranoic tendencies. Comparison indicates that all three parameters characteristic of definite schizophrenics are definitely larger than those of others, in agreement with theoretical expectations.

A COMMITTEE REPORT ON SCHIZOPHRENIC LANGUAGE

by John R. Parks

Monsanto Chemical Company, St. Louis, Missouri

In June, 1958, a committee composed of three scientists of the Monsanto Chemical Company and three psychiatrists of St. Louis State Hospital was voluntarily formed to study the possibilities of information theory and computer techniques in advancing understanding of schizophrenic language.[1] This report is a brief resume of achievement in the eight evening meetings to date.

From the beginning it appeared that statistical analysis of rank-frequency data of long messages from various sources, normal and abnormal, was the foundation of the work of the committee. Zipf and Whitehorn have discussed rank-frequency data obtained from a very wide variety of sources, some of which were schizophrenics. However, the discussions, though brilliant, were hampered by lack of numerical analysis of the data and hence were cast in a semi-qualitative mold (Whitehorn & Zipf, 1943; Zipf, 1949). Rank-frequency data consist of counts of occurrence of unique words in messages of 15,000 or more consecutive words, with counts ranked according to size, beginning with the greatest count as rank one. Plots of such data on log-log paper were analyzed by Zipf to come to some conclusions concerning normal and abnormal language.

The log-log plots of data obtained from counts of Joyce's *Ulysses* and of a large composite of newspaper English from the *New York Times* suggested an hypothesis to Zipf and Whitehorn which is now of historical significance only. They postulated that individuals will minimize their effort in the generation of long messages, and they used a pseudomathematical method to derive from the hypothesis an equation which accounts for the shape of the log-log plots of the *Ulysses* and *New York Times* data. Zipf's law was the result of this study. To the writer's knowledge, this is the first scientific

[1] The other members of the committee in addition to the author are: A. K. Busch, M.D., Clinical Director, and A. Heusler, M.D., Psychiatrist, St. Louis State Hospital, St. Louis, Mo.; G. Ulett, M.D., Director, Malcolm Bliss Hospital, St. Louis, Mo.; Leon Cooper, Ph.D., Manager, Applied Mathematics Section, and T. J. Williams, Ph.D., Group Leader, Systems Section, Research and Engineering Division, Monsanto Chemical Co., St. Louis, Mo. The committee wishes to acknowledge the help of G. C. Harris and W. A. Hauser in the preparation of the Jackson M. data.

attempt to account for the rank-frequency data obtainable from long messages.

The effect of the size of the vocabulary is removed by dividing the counts by the total number of words per message. This converts rank-frequency data to rank-probability of occurrence data. At high ranks, groups of words with the same probability of occurrence appear and each group is assigned an average rank which is the average of the lowest and highest ranks in the group; the result is a final conversion of rank-frequency data to a more statistically usable form, namely average rank-probability of occurrence data.

B. Mandelbrot (1953) published a thesis in which he proposed a rational explanation of the remarkable form and stability of the average rank-probability of occurrence data. He examined the problem of optimizing the generation of long messages by a message generator loading a cable with signals, one of which denotes a space or silence. He defined the process of optimization as one in which the message generator will use unique words with such a frequency that the rate of transmission of engineering information is a maximum. The argument involved the application of the calculus of variation to a combination of three fundamental aspects of the problem; first, that the probability of a message is a decreasing function of the length of the message; second, that the variety of messages of the same length is an increasing function of the length; and third, that the meaning of the words or the messages is not important in the solution of the problem and, therefore, may be neglected. The final result of Mandelbrot's work was equation (1).

$$P = P_o (\bar{R} + B_o)^{-\alpha} \qquad (1)$$

where

P = probability of occurrence of words
$\bar{R}$ = average rank of the words
P_o, B_o, α = parameters; the values of which determine the function.

Equation (1) was rigorously derived from the definition of engineering information and the properties of real networks in the transmission of it. The term "engineering information" is used in an attempt to circumvent confusion with the usual notion of the term "information" which is embedded in semantic overtones and value judgements. If one knows the signals and the constraints under which the message generator uses them, the rate at which the generator produces the signals, and the rate at which the cable transmits the signals, one could calculate the parameters P_o, B_o, and α by Mandelbrot's theory, thus specifying in advance the average rank-probability of occurrence of unique words in any long message produced by the generator and cable combination. Equation (1) is general and contains Zipf's law as a special case.

Mandelbrot's theory may be set aside, possibly temporarily, and equation (1) viewed only as a correlation equation for analysis of rank-frequency data; in which case values of the parameters P_o, B_o and α would be so computed that a best fit of equation (1) to the data at hand would result. "Best fit" is used in the statistical sense that the sum of the squares of the deviations of computed values from the given data is a minimum. A. W. Dickinson, of the Applied Mathematics Section of the Research and Engineering Division of Monsanto, has developed a general IBM 704 nonlinear regression program which has been used to obtain the values of the Mandelbrot constants for nine sets of data. Table 1 shows the results of the machine computation. The

TABLE 1

COMPARISON OF CHARACTERISTIC CONSTANTS IN MANDELBROT'S EQUATION OBTAINED FROM VARIOUS NATURAL INFORMATION SOURCES

		P_o	B_o	α
1	Joyce, *Ulysses*	0.1138		1.026
2	5 yr. old girl*	0.1868	1.120	1.050
3	Jackson M., Vol. 1	0.1701	1.547	1.038
4	Jackson M., Vol. 3	0.1939	1.614	1.071
5	Standard English	0.1786	1.6542	1.0927
6	Charles M.*	0.2830	2.9139	1.1442
7	Harold D.	0.5278	4.563	1.265
8	Childish Adult*	0.3909	5.1320	1.1921
9	Helen B.*	0.5365	5.3768	1.2005

* Data read from graphs in Whitehorn and Zipf (1943).

standard deviations of the fit of equation (1) to the data varied from 5 to 15 per cent and the coefficients of correlation were not less than 0.995 in all cases, thus indicating excellent fit. The less accurate fits were for data obtained from graphs as indicated in the Table.

The results are arranged in Table 1 in increasing order of the value of B_o. The value of B_o for Joyce's *Ulysses* is not given because the data available at the time of the calculation did not include the frequencies of words with ranks less than 10. However, it is to be expected that the value would be least if the data were available. Mandelbrot's theory indicates that the fewer the constraints on the generation of messages, the smaller the value of B_o and in *Ulysses*, James Joyce showed signs of ignoring the constraints that are normal to standard English. The theory also implies that B_o cannot be less than 0.5. The largest value of B_o in Table 1 is associated with the language of Helen B. who was diagnosed as schizophrenic.

Studies of the language of schizophrenes from the point of view of grammar and logic indicate that such people do not use the language to its fullest capacity, thus implying that B_o for schizophrenes should be greater than that for standard English as is shown in Table 1. (See also Woods, 1938.) Harold D. and Charles M. are schizophrenics. Standard English appears to divide the group into psychotic and nonpsychotic subgroups. The language of the childish adult places her with the schizophrenics. There is evidence that Jackson M. wrote the three volumes of *Right-wrong* before he was hospitalized as schizophrenic (Thale, Hauser, & Harris, 1959). More cases must be studied in order to substantiate the implication in Table 1.

The data for James Joyce's *Ulysses* were obtained from a table of rank-frequency data in Zipf's book *Human Behavior and the Principle of Least Effort* (1949). This table contains about 15 entries with the words of rank less than 10 not listed, and represents a count of 260,430 words.

Joyce lived from 1882 to 1941. At the University College in Dublin, he specialized in modern languages. He was a gifted tenor and amassed a vast knowledge of music. He began the study of medicine, considered founding a great Irish daily, opened a motion picture theatre in Dublin, and engaged in the legitimate theatre in Zurich. This was in addition to his literary work. This wide range of interests may serve to explain the rather extraordinary combination of tradition and experiment in his writings. It is possible he was familiar with the early statistical studies of language and the foundations of modern psychology. It must be remembered that *Ulysses* marked a transition from the eminently readable *A Portrait of the Artist as a Young Man* to the so-called unreadable *Finnegans Wake*—unreadable chiefly because it offers the reader a range of choices of meanings rather than just the meaning expressed by the author.

Jackson M. was a citizen of St. Louis who was moved to write a three volume work which he collectively entitled *Right-wrong*. The material appears to be autobiographical. The style is bizarre in many sections and the author seems quite involved in his relations with a girl and the females in his family. It is not known how long after writing *Right-wrong* he entered the hospital.

Early last year Messrs. Hauser and Harris, two students of the St. Louis University Medical School, undertook the listing of 15,000 consecutive words from the initial chapters of the first and third volumes of *Right-wrong*. These lists were keypunched on IBM cards so that they could be sorted and a rank-frequency tabulation made.

The 10,000 word message Zipf analyzes for the case of the five-year old girl was the transcript of a set of recordings made by a Dr. R. S. Uhrbrock about 1935 when he was doing some research in education. Zipf was interested in this case because the speech of psychotics if often likened to the speech of children. The data we have used to obtain the Mandelbrodt constants were read as accurately as possible from a log frequency versus log rank graph published by Whitehorn and Zipf (1943).

The Standard English data are from a study of 100,000 words forming a composite of newspaper writing, personal letters,

speeches, religious literature, etc. (Dewey, 1923). These data did not contain a listing of words occurring 10 times each and less.

Charles M. was a patient before the Civil War at the McLean Hospital for the insane in Boston. He left a rather voluminous set of letters and autobiographical material. Whitehorn and Zipf (1943) made a count of 47,850 running words of this material which they published as a log-log graph. Our data were read from this curve.

The childish adult was a 33-year old woman who was a highly trained chemist and a productive scientist. However, she was childish in the character and expression of her emotional relationships. She also was a patient of the McLean Hospital. After her discharge from the hospital, she wrote frequent letters to her favorite physician in which she asked permission to call him "Dear Daddy" in her letters. This long distance therapy continued for six or seven years. She continued her work in research from which she gained considerable scientific distinction, including an honorary Doctor of Science. It was from her letters to the physician that Zipf obtained frequency-rank data on 50,000 consecutive words. He published the results as a log-log graph which we used to obtain our data (Whitehorn & Zipf, 1943).

Helen B. was a 23-year old unmarried college student who entered McLean Hospital after screaming denunciations of her father, mother, and sister and threats to kill them. She was diagnosed as a schizophrene with paranoiac tendencies. During her hospitalization, she wrote sufficient letters for a statistical analysis of 50,000 running words. Zipf published the data as a graph from which we got our data (Whitehorn & Zipf, 1943).

The rank-frequency data for Harold D. were obtained from a transcript of recordings of interviews by a psychiatrist at the Metropolitan State Hospital at Waltham, Massachusetts. Dr. James G. Miller, Director of the Mental Health Research Institute, The University of Michigan, was kind enough to lend to the committee for analysis the tabulation of the data done by him and Mr. Russell F. W. Smith. The tabulation covered 10,880 running words.

The most accurate analyses we have in Table 1 are those for Jackson M., Standard English, and Harold D. The *Ulysses* data were not extensive enough since they did not include low rank data; the others are quite inaccurate because the data were read from graphs, as has been indicated.

Since Mandelbrot's theory appeared, there has been considerable discussion of the assumption that the individual seeks to maximize the rate of generation of engineering information by his choice of words in the generation of long messages. It is felt that this is too strong an assumption and possibly an unnecessary one. In 1957, G. A. Miller displayed an example of a generator and a cable for which he could derive the Mandelbrot equation [equation (1)] without the use of the maximizing assumption. The generator was a monkey and the cable a typewriter. The monkey typed long messages under the following constraints: the probability of the monkey touching the space bar is specified as P_s; the probability of the monkey touching a letter key is distributed equally among the keys on the typewriter and is $(1\text{-}P_s)/A$, where A is the number of the letter keys on the typewriter; the monkey was further constrained to no consecutive touching of the space bar. Using these assumptions, Miller derived a form of Mandelbrot's equation as shown in equation (2).

$$P = \frac{P_s}{1-P_s}\left\{\left[\frac{A+1}{2(A-1)}\right] \cdot \left[R + \frac{A+1}{2(A-1)}\right]\right\}^{-\log \frac{A}{1-P_s}/\log A} \tag{2}$$

Here it is seen that the Mandelbrot constants P_o, B_o, and α are expressible in terms of P_s and A.

At this time the committee has no quarrel with Mandelbrot's or Miller's assumptions and is only seeking reasonable hypotheses on how the central nervous system may operate in the generation of messages, thus throwing some additional light on schizophrenia. To this end, G. A. Miller's solution of the monkey and typewriter problem has

been modified to include not only the probability of a space, P_s, and the number of letters, A, the monkey may choose freely, but also T_1, the time it takes the monkey to choose a letter, and T_2, the response time per character of the typewriter (Miller, 1957). The solution of this problem yields equation (1) in such a form that

$$P_s = \frac{P_o}{P_o + B_o^{-\alpha}}, \qquad A = \frac{2B_o + 1}{2B_o - 1},$$
$$\frac{T_2}{T_1} = \alpha \frac{\log A}{\log \frac{A}{1 - P_s}} \qquad (3)$$

Equations (3) permit the specification of random word generators the messages of which will have the same rank-frequency statistics as the messages from natural message sources, such as shown in Table 1. In the consideration of these equations, the committee is exploring the idea that events occur in the central nervous system on two major time scales: one, T_2, corresponding to the capacity of the central nervous system to transmit engineering information in the form of neural pulses; and the other, T_1, corresponding to the initiation of the pulses to generate messages. The committee is the first to admit that this model is extremely simple and probably useless. However, we feel that, considering the value of the objective, it is worthwhile to determine whether it is a foot-hold, toe-hold, or no hold at all as a guide to future work.

This report is a rather bare outline of the work the committee has done to date on the statistics of the language of schizophrenes. The major contribution is the accurate determination of the Mandelbrot constants for nine different types of real message generators, pending the elucidation of the physical significance of the constants. The committee has begun a very recent mathematical study of another approach to the statistics of language suggested by H. A. Simon (1955; 1960).

Since the writing of this report, the complete word statistics of Joyce's *Ulysses* has been analyzed (Hanley, 1937). The values found for the Mandelbrot constants are as follows: $P_o = 0.1302$, $B_o = 1.285$, $\alpha = 1.044$. It is to be noted that the contention is correct that in the case of *Ulysses*, B_o could not be zero as suggested by Zipf. However, this value of B_o is not the least; the 5 year old girl has a smaller B_o.

Analysis of some samples from Joyce's *Finnegans Wake* is in progress and will be reported on at a later date.

The committee urges anyone who has vocabulary data on long messages from sources with well-defined backgrounds, or any other pertinent data, to contact them so that the project may be pushed forward more rapidly.

REFERENCES

Dewey, G. *Relative frequency of English speech sounds.* Cambridge: Harvard University Press, 1923.

Hanley, M. L. *Word index to James Joyce's Ulysses.* Madison: Univ. of Wisconsin Press, 1937.

Mandelbrot, B. *Logique, langage et theory de l'information.* University of Paris: Publ. Inst. Statist. 2, 1953. (Thesis)

Miller, G. A. Some effects of intermittent silence. *Am. J. Psychol.*, 1957, 70, 311–314.

Simon, H. A. On a class of skew distribution functions. *Biometrika*, 1955, 42, 425–440.

Simon, H. A. Some further notes on a class of skew distribution functions. *Inform. & Control*, 1960, 3, 80–88.

Thale, T., Hauser, W. A., & Harris, G. C. Schizophrenic writing, a clinical study. 1959, to be published.

Whitehorn, J. C., & Zipf, G. K. Schizophrenic language. *Arch. Neurol. Psychiat.*, 1933, 49, 831.

Woods, W. L. Language study in schizophrenia. *J. Nerv. Ment. Dis.*, 1938, 87, 299.

Zipf, G. K. *Human behavior and the principle of least effort.* Reading, Mass.: Addison-Wesley, 1949.

(Manuscript received March 15, 1960.)

BOOK REVIEWS

Franz Alexander, M.D., *The western mind in transition.* New York: Random House, 1960.

DRAWING on his personal experience as a practicing psychoanalyst and astute observer of the human scene, Dr. Alexander delineates in this book the radical changes which have been taking place in Western culture in the past fifty years.

Born in an era when the search for absolutes in knowledge, taste, and morals still predominated the mores of the time, and in a culture where knowledge for its own sake was regarded as the highest form of endeavor, he traces the end of this era and the beginnings of the relativistic approach in science, art, politics, and national economies. While the first world war is generally regarded as the event which precipitated this change, the precursors of this transition had already been active in a number of disciplines: the philosophies of Kierkegaard and Nietzsche, the post-impressionist paintings of Van Gogh and Cézanne, the writings of Shaw and Ibsen, the physics of Einstein and Planck, and the psychoanalytic theories of Freud and his followers. The all-pervasive nature of this change gives it the force of a major historical trend, the responsibility for which cannot be assigned to any single cause.

Whatever its origins, the scientific revolution has entered the social sphere and the statistical concepts of the atomic physicist apply with equal validity to our dealings with one another. The individual is now judged not by his intrinsic and unique qualities but by the more easily quantifiable parameters of height, age, salary, examination grades, and other statistically derived measures. The growing threat that this process poses to man's individuality has been recognized not only by the psychoanalysts under the term "loss of ego identity" but is also the target of the existentialist school of philosophers and writers. They too have pointed out that the estrangement of man from nature and from himself has left him with a feeling of helplessness and despair. The most common form of escape from this intolerable situation, that of identifying oneself with some firmly established organization, has resulted in a wholesale increase in conformity with a consequent loss of spontaneity and creativity.

The problem posed by Dr. Alexander is how to restore creativity and remove man from the ant-like existence enforced by the increasing trend toward greater and greater organization. Here his training in psychoanalysis suggests the solution. A child develops his feeling of identity, his ego, by groping experimentation. By this process of play, and depending on his own resources, a child can freely develop his own uniqueness within the larger framework of existing traditions. If man is to make creative use of the increased leisure brought about by mechanization of production, he must apply the same principles of play to his life. Creative change, according to Dr. Alexander, can only come from the not immediately useful realization of individual strivings, curiosities —sometimes even oddities—which appear idle, playful and non-utilitarian at the moment. In support of this thesis, he cites examples of successful scientific innovation and likens the activity of the scientists engaged in such work to a form of uninhibited play. The same spontaneous learning process which is the essence of play is also operating in the discovery of new scientific principles. This is to be contrasted with the widespread tendency in our present educational system of relying on the process of identification wherein the students are presented with the ready-made results of previous learning. While this may be economical, something is lost by this economy, since the knowledge acquired by repeating the original steps of discovery using one's own efforts results in a more basic understanding of the problems involved.

Dr. Alexander exhibits a deep understanding of the problems of the creative

scientist, obviously from first-hand experience. He writes very movingly on the need for basic research and the atmosphere of freedom and lack of constraint under which such research can best flourish. His discussion on the role of the scientist in society should be required reading for administrators and public officials with responsibilities for promoting the growth of science.

The author naturally devotes some space to his own science of psychoanalysis, its early history and the attraction which it offered to him as a perfect blend of the humanistic and scientific approaches toward life. There is also considerable discussion of his position on the integration of psychoanalysis into the medical curriculum. He regards the separation of psychoanalytic training from the rest of medical and psychiatric instruction as a residue of the past, due largely to the reactions of Freud and his early followers to criticism and hostility from the rest of the medical fraternity. The time is now ripe, according to Dr. Alexander, to teach psychoanalysis in the traditional places of medical training, the universities and teaching hospitals.

This is a highly stimulating book. Written, as it is, from the personal point of view, it enables the reader to view the changes taking place in our culture through the eyes and mind of one of its most careful observers, and this reviewer, at least, has found it a very rewarding experience.

WILLIAM J. HORVATH
Mental Health Research Institute
The University of Michigan

Robert M. Featherstone and Alexander Simon (Eds.), *A pharmacologic approach to the study of mind.* Springfield, Ill.: Charles C Thomas, 1959.

THIS is a printed report of a three day symposium with the same title held in January, 1959, under the auspices of the San Francisco Medical Center of the University of California and the Langley Porter Neuropsychiatric Institute, with financial support from Lakeside Laboratories, Inc. of Milwaukee, Wisconsin. There were 41 participants in the program plus a dinner speech by Aldous Huxley entitled "The final revolution," in which he talked of the application of technology to governmental control over human conduct fantasied about in *Brave New World.* Huxley fears that this kind of world is coming closer much faster than he expected, and wants scientists to abandon from time to time their narrow specialization to worry together about this trend, and to try to seek ways of mitigating or averting the loss of individual freedom which he fears will result.

The first half of the symposium was devoted to papers which are essentially progress reports by specialists (in pharmacology, psychiatry, neurochemistry, psychology, and statistics) striving to keep one another informed about a rapidly expanding field. There are many items of information but no major insights, though this may well be representative of the state of the art in psychopharmacology. Among the presentations, the paper by Zeller on the concept of enzymes as an interpretative principle in physiology, pathology, and pharmacology presented material not frequently discussed elsewhere. The papers were followed by discussions which were particularly lively.

The second half of the symposium was devoted to reports on tranquilizers, hallucinogens, stimulants (monamine oxidase inhibitors), and psychic energizers.

In evaluating this publication, one could use two standards: other symposia, or books by one or more authors. While this publication certainly cannot afford comparison with the latter standard, it measures up well against other symposia. Though it gives the flavor of a real discussion, the individual papers are carefully written and contain less duplication of topic than is often the case.

STEVEN G. VANDENBERG
University of Louisville

T. E. Ivall, *Electronic computers. Principles and applications.* London: Iliffe & Sons Ltd., 1956; New York: Philosophical Library, 2nd ed., 1960.

THERE are by now quite a few books about electronic computers and it

becomes on the one hand easier to review such a book because a comparison is possible with other works, and at the same time more difficult because a reviewer finds it often difficult to read yet another in its entirety.

Recently introduced into this country by the American publisher, the present volume was written for the intelligent British layman who wants more than a vague understanding of the principles of computer circuitry. The subtitle is somewhat misleading because there is not a word about programming a problem for a digital computer so that the reader will have no idea about the many man hours of work involved in preparing computer programs, nor about the complete dependence of the machine on such man-made programs. In this respect it is strange to find that the book is aimed also at businessmen who may be prospective buyers of a computing system.

This crisply written book contains simple but accurate descriptions of some of the types of circuits used in computers and contains many photographs of parts of computers. Another somewhat unusual feature is the considerable space devoted to a discussion of analogue computers. Actually the discussion of analogue type circuitry lends easily to the somewhat more difficult topic of digital circuitry.

The first edition was written by seven authors in addition to Ivall. This second edition has been largely rewritten by Mr. Ivall to bring the book up-to-date on recent developments. Little of the original contributors' work remains though "their inspiration is still behind the book."

STEVEN G. VANDENBERG
University of Louisville

Charles A. Beard, asked if he could summarize the lessons of history in a short book, said he could do it in four sentences:

1. Whom the gods would destroy, they first make mad with power.
2. The mills of God grind slowly, but they grind exceedingly small.
3. The bee fertilizes the flower it robs.
4. When it is dark enough, you can see the stars.

HARRY H. GOODE (1909-1960)

Harry H. Goode, Professor of Electrical Engineering and Chairman of the Management Science Committee at The University of Michigan, died in an automobile accident on the morning of October 30, 1960. His loss will be deeply felt by all who knew him through his teaching, his frequent lecture appearances, his many publications, his work in professional societies, his consulting activities, his stimulating participation in conferences, or directly through his warm friendship.

His many published papers touched upon statistics, simulation and modeling, vehicular traffic control, and system design. His major published work is the book *System Engineering*, of which he was senior author with R. E. Machol. The book was the outgrowth of the very successful and valuable course which he introduced at The University of Michigan under the title, "Large Scale System Design."

Professor Goode was known internationally for his work in management science, system engineering, and electronic computers. He was Chairman of the National Joint Computer Committee of the Institute of Radio Engineers, American Institute of Electrical Engineers, and Association for Computer Machinery. As its representative to the international group, he played an important part in the formulation of the charter under which these three organizations will join to form the American Federation of Information Processing Societies, which in turn will represent this country in the International Federation of Information Processing Societies. His broad experience with computers and his experience in national computer functions led to his participation as one of the group of eight Americans who visited Soviet computer establishments in 1959. More recently, he visited Turkey on an advisory mission for NATO.

Our profession has lost one of its outstanding members—a man of rare versatility, talent, vigor, and vision.

MERRILL M. FLOOD
Mental Health Research Institute
The University of Michigan

COMPUTERS IN BEHAVIORAL SCIENCE

Publication of this department is partially supported by a grant from the National Science Foundation to the chairman of its editorial committee, Steven G. Vandenberg

Some Problems and Approaches to Automation of Medical Diagnosis,*

Philip A. Smith, System Development Corporation, Santa Monica, California

INTRODUCTION

A common misconception is that a medical diagnosis represents hard fact rather than a physician's professional judgment, estimation, and opinion (Bostian, Smith, Lasky, Hover, & Ging, 1959). Obtaining a medical history and recording physical symptoms is a complex judgmental process rather than a simple enumeration and tabulation of discrete, manifest, informational items. There has been little study of the consistency (reliability) of a physician's judgments over time, but evidence from the psychological area suggests that reliability decreases as the complexity and difficulty of discrimination increases (Meehl, 1954). Inspection of hospital records shows that when different physicians have occasion to see a patient independently, they sometimes tend to elicit dissimilar medical histories. Clinical experience shows also that patients differ markedly in the degree to which they feel free to talk or are able to describe their symptoms. Thus, the determination of what represents factual raw data for diagnosis is itself a formidable problem. This fundamental consideration must be recognized in any approach to automating medical diagnosis.

OBJECTIVE MEASURES

It can be easily understood, then, that there is a temptation and a need for proponents of automation to seek objective measures of patient function—measures such as blood pressure, temperature, electrocardiograms, X-rays, electroencephalograms, laboratory tests, and so on. However, the use of laboratory tests and instrumented data-gathering procedures may not add much in the way of reliable information. For example, the basal metabolic rate (BMR) has been used extensively as one measure of functioning, but it is considered by some medical experts to be nearly valueless as a diagnostic aid. While the procedure is objective, the variation in individuals from day to day tends to invalidate the results of any single measure. Similar objections have been lodged against the use of the electroencephalogram (EEG). Not only is there variability in a single individual's EEG record over time, but the reliability of independent EEG readers or interpreters is poor. And the few normative studies of EEG's made to date suggest that the percentage of so-called "abnormal" records may range as high as 40% in a "normal" population! In short, the precision of data on which diagnosis is based, even if those data are gathered by relatively objective instrumented procedures, is often more apparent than real. But to say that such measures are less than ideal does not mean that we should no longer attempt to use them. One can raise this question—for example in connection with the EEG—is the *procedure* unreliable or is it our lack of broad normative data that cuts down diagnostic accuracy?

One of the most promising approaches to problems in this area has been stimulated by the recent "man in space" and "animal in space" programs. Here, the job calls for sturdy, compact, lightweight instruments to be attached to an organism enclosed in a space capsule; to monitor and to record or

* Paper presented at the Western Psychological Association symposium, "The future role of computers in diagnosis." San Jose, California, April, 1960.

telemeter back to earth a continuous series of physiological measures. A number of medical electronic firms are developing microminiature sensory buttons, transducers, and instrumented suits or vests for this purpose. One firm is developing endoradiosonds, or "radio pills," which may be swallowed and which will transmit for a short time indications of gastric acidity, pressure, and mobility. The Air Force (ARDC, 1960) has reported acceptance of a physiological data acquisition system which will continuously monitor and telemeter a dozen bodily measures, including heart and respiration rate, blood pressure, and galvanic skin response (GSR).

It may soon be possible, with little discomfort or even awareness on the part of a subject, to obtain continuous, round-the-clock measures of his physiological response to the stresses and strains of daily life. We could get a good picture of the extremes and ranges of individual variability and gather perhaps startling normative data. Those who have struggled with the difficulties of getting reliable GSR's in an experimental situation, or who have considered the imposing problem of obtaining continuous stress variable measures on a patient during a psychotherapeutic hour, must recognize the tremendous technological advances implied. Of course, computer processing of such prodigious quantities of monitored data is necessary.

These are the advances which have given rise to the Sunday Supplement stories about such things as diagnosis cubicles—instrumented examination booths with a direct line to a medical computer—and automated hospitals where a computer keeps watch over patients lying in instrumented beds.

SUBJECTIVE DATA

Now, what about diagnostic measures that are not strictly physiological? Let us consider other, more subjective classes of diagnostic data. There are estimates that 50 to 80 percent of all medical problems include elements of psychological health: some of the symptoms of illness are in part the equivalents or sequelae of anxiety. The difficulty in making valid inferences of psychological states is well known. What can be done about these subtle variables, that so far as we know cannot be routed through any kind of sensing device except another human being? One approach is to route clinical observation and interview data through a judge. Assuming judges agree, the data can then be processed by a computer (Bostian, et al., 1959). Or, the diagnostic middleman can in some cases be skipped: studies made of clinical impressions as complex as diagnosis show that a simplified procedure, such as employing a check list of signs or indicators which are checked against a statistical norm, may equal or exceed the efficiency of practiced clinicians making their most artistic and intuitive judgments (Kelly, 1957; Meehl, 1954).[1]

This check list approach has led to interest in automating the interpretation of diagnostic inventories such as the familiar Cornell Medical Index and the MMPI. It would not be difficult to display test items on a cathode ray tube, provide a push button response mechanism, rapidly record and compare a subject's data profile against a central library of thousands of profiles, and then compute one or more provisional, general diagnoses.

Clerks and psychometricians have been doing this and more, manually, for years. To do it by machine is more exciting, much faster, and a great deal more expensive. It is a tentative, somewhat primitive first step. But it is coming. And I think psychologists and physicians have a responsibility to help guide and direct these efforts, for they quickly recognize the limitations in such approaches: the entire procedure is only as good as the test items or other data on which it is based; it is dependent on the ability and willingness of the subject to read, to comprehend, and to respond; it is dependent on the breadth and extent of the comparison data, and on the very meaning and utility of diagnosis itself.

[1] When E. L. Kelly (1957) examined the criteria by which candidates for medical school were selected, he found that decisions based on a fairly simple statistical index excelled the decisions of medical faculty who evaluated the "total candidate." The parallel is easily extended to the diagnostic process, by which the physician attempts to evaluate the "total patient."

OTHER APPLICATIONS

There are a number of other areas, apart from diagnosis, in which the computer may assist as a control, bookkeeping, or research device. I will mention only in passing that there is some thinking about much more sophisticated use of a computer—to emulate the cognitive processes of a medical diagnostician, rather than serve as a rapid but non-intelligent statistical slave (Vandenberg, 1960). Other proposed computer applications range from a suggestion to use a small digital device to control the administration of anesthesia (Grabbe, 1959), to a plan for a world-wide public health information processing system to predict and monitor the incidence and spread of disease. However, the first attempts at automation call for setting up the machinery to handle hospital records and public health data on a more limited basis; for scientific use of computers in large-scale research projects; and in pilot systems for direct multivariate data collection, conversion, diagnostic analysis, and display to the individual physician (Ledley & Lusted, 1959; Roach, Pearce, Moore, Smith, & Walsh, 1959). I will comment on two other applications: first, in the area of drugs and pharmaceuticals; and second, hospital administration.

There may be ready use and acceptance for rapid-access computer-stored catalogues of pharmaceutical information, such as drug composition, cross-indexed lists of variants and derivatives, trade names, dosages, side effects, substitutes, and so on. In the tranquilizing field alone, new compounds are produced each year literally by the thousands, so the need for good indexing procedures is likely to grow. I understand that the Bureau of Standards is already making some effort in this direction.

Full clinical trials of the usefulness of promising drugs is a separate problem, with difficulties similar to those inherent in any study of treatment outcome. In the field of tranquilizing and energizing drugs, which are now used extensively both in psychiatry and in general medical practice, there already exists machinery for full-scale clinical trials with modern data processing methods to aid in analysis and evaluation of results.[2]

A second area in which the computer may assist is in hospital administration. In a typical general medical and surgical (GM&S) hospital, peak work loads occur in certain wards at fairly predictable intervals. Staff shortages, especially in the area of nursing, have led some hospitals to shuffle staffing schedules in order to provide personnel where and when they are most needed. Similarly, in apportioning operating funds, hospital administrators attempt to juggle their spending to meet essential needs. Within total hospital systems, for example the Armed Forces and the VA, there exist problems of even greater magnitude: which hospitals can be closed, where should new ones be built, and how should they be staffed to meet a growing, shifting population? How may operating funds for the total system be budgeted and disbursed? These problems are essentially linear programming functions and may eventually be solved by computer techniques.

The advantages of cataloguing and processing patient information in total hospital systems such as these are obvious. Especially in the Veterans Administration, where the patient population tends to be a recurring but constantly aging one, will efficient cataloguing and storage of personal data and diagnostic information for longitudinal study be immensely useful.

Now, one final caution which applies whether one seeks diagnostic research data, gathers broad public health information, or attempts to standardize the information entered on hospital records. When one engages in the collection and storage of basic normative data, or searches for common factors, then common measures on a large group of subjects must be obtained. Patients sometimes distort or attach great personal significance to portions of their medical examination. May nonvolunteers, for ex-

[2] The Veterans Administration Central Neuropsychiatric Research Unit, Perry Point, Maryland, has for several years been carrying out such studies on a co-operative nationwide basis, and at present has underway a program of research to evaluate routinely new drugs which hold promise for psychiatry.

ample random admissions to a hospital, be given massive examination procedures, some of them frightening, painful, or even moderately dangerous, for the sake of gathering a common pool of information on a large group of subjects? This is a practical or perhaps an ethical problem which intrudes on the diagnostic task. This objection can be circumvented by the use of volunteer subjects, or by coercion of prisoners or civil service job applicants to serve as subjects. Of course, one can badly bias his sample by using primarily such strata of *S*'s. Fortunately, this is another of those problems that psychologists are well equipped to recognize and to handle.

REFERENCES

ARDC News Release No. 60-60-R, March, 1960, Holloman Air Force Base, New Mexico.

Bostian, D. W., Smith, P. A., Lasky, J. J., Hover, G. L., & Ging, R. J. Empirical observations on mental-status examination. *A. M.A. Arch. Gen. Psychiat.*, 1959, 1, 253–262.

Grabbe, E. M. Digital computers for automatic control in medical applications. Paper presented at the Second Annual Conceptual Symposium, Automedic Research Center Inc., Montgomery, Alabama, March, 1959.

Kelly, E. L. Multiple criteria of medical education and their implications for selection. In *Report of the fourth teaching institute*. Evanston, Ill.: Association of American Medical Colleges, 1957.

Ledley, R. S., & Lusted, L. B. The use of electronic computers to aid in medical diagnosis. *Proceedings of the IRW: Bio-medical electronics issue*, November, 1959, 1970–1977.

Meehl, P. E. *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota Press, 1954.

Roach, C. J., Pearce, M. L., Moore, F. J., Smith, E. C., & Walsh, J. E. *SP-102, Data processing in the medical sciences*. Santa Monica, California: System Development Corporation, October, 1959.

Vandenberg, S. G. Medical diagnosis by computer: Recent attempts and outlook for the future. *Behav. Sci.*, 1960, 5, 170–187.

(Manuscript received September 1, 1960.)

Suggestions for Self-Adapting Computer Models of Brain Functions,

Leonard Uhr and Charles Vossler,[1] *Mental Health Research Institute, The University of Michigan; and System Development Corporation*

All computer programs that model functions characterized by terms such as "learning," "problem solving," "concept formation," "perception," have been developed to cope with specific problem spaces, for which specific solution transformations were known to exist. (See, for example, Newell, Shaw, & Simon, 1960; Gelernter, 1960; Samuel, 1959; Friedberg, 1958; Kilburn, Grimsdale & Sumner, 1960.) These programs were written to possess operators which we know can arrive at the solution. The chief interest in them probably lies in the various methods they use for assessing the problem space, paring down overlarge spaces to manageable size, and directing search as a function of information gathered. These programs are, essentially, embodiments in an information-flow model of procedures for transforming input information that the programmer knows will be presented to the computer into a desired output form that is known to exist and to be achievable by operators that the computer has been given in advance.

The following is a suggestion for a general purpose computer (or brain model) that is *not* designed for any specific input-operator-output space; but rather will learn to adapt and compose itself into an appropriate model for operating upon any space presented to it. It is suggested that this model may reach the limit of generality that we can expect of any intelligent computer (including the human brain). First, some results are reported for a pattern recognition program that uses this approach, then an approach to

[1] We would like to thank William Karush, Center for Research in System Sciences, System Development Corporation, along with many other members of the Center, for the opportunity to discuss and work on some of these problems. We would also like to thank John Gyr, John Dwyer, Al Cafagna, and others at the Mental Health Research Institute for much stimulating discussion.

language learning by computers is outlined. Finally, a very simple basic unit, one that appears to be plausible from a neurophysiological stand-point, and easily realized with logical circuitry of the sort used in computers, will be described.

A PATTERN RECOGNITION SIMULATION ON A DIGITAL COMPUTER

A brief sketch of a computer program for pattern recognition of visual forms in which some of these principles are embodied follows. We are now in the process of testing and extending a pattern recognition program, for the IBM 709, that is supposed to learn to recognize patterns such as alphanumeric letters and line drawings of simple objects. In this program, patterns are initially presented to a computer without any operators or any output space of concepts about meanings of output sets. Operators are generated by two methods: (a) by extracting (if you will, "imitating") fragments of the inputs-as-known to the computer via its input transducing elements, and (b) by generating at random Boolean functions, or "n-tuples" of these input element states (this is an extension of Bledsoe and Browning's [1959] generation method, to achieve an optimum set of sufficiently powerful operators). These operators then map inputs into output sets, and feedback as to the appropriateness of the mapping controls generation and refinement of subsequent operators.

The input space (in this program, a 20×20 0–1 matrix, giving a space of 2^{400} states) upon which pattern recognition operators must act is almost certainly too large for the above method to work. Therefore, at present, two psychologically plausible constraints have been placed upon the random generation of operators within this space, in order to cut it down to size. These have been (a) generation of operators that are functions of *connected* elements only, and (b) generation of these operators within the space of a *smaller* matrix, and *translation* of this matrix across the larger space. These restrictions make the operators equivalent to simple neural nets that are repeated in parallel. We are now attempting to replace these restrictions by "natural" partitionings of the problem into a series of problems, in which the first set of operators will map inputs into a set of sets of output operators (which have been classified by a set of operators that has learned to recognize the patterns among operators and the patterns across simple transformations).

The first two machine runs gave 80% and 90% success in recognition across seven quite variable examples of each of five patterns—this the first time through, before the places at which adaptive learning occurs had had much chance to reach an optimal adjustment. Of special interest is the way the various particular amplifiers of specific operators and the characteristics of their outputs, originally set at extremely low values, so that they would have virtually no effect until they gave evidence that they would be successful, were gradually being tuned up as a function of this evidence. Similarly, several new operators had already been generated (by extraction—or "imitation"—of a randomly chosen portion of a unknown input) to replace operators that had been judged useless. Finally, in the last machine run to date, the program began without *any* operators, but merely the rule to generate five new operators by extraction from each unknown input, up to a limit of 40. Again, within the period of a short test run, the program's processing abilities had improved from 0% to better than 90%.

In passing, note that most pattern recognition programs (and, in fact, most conjectures as to models for form perception, of which present-day computer programs are, in the judgment of the authors, among the most interesting) are, when examined from this point of view, simply arbitrary selections of sets or types of sets of operators. For example, Rosenblatt's Perceptron (1958) operates with one set of randomly chosen single element sets from the entire input set. Selfridge's Pandemonium (1960) operates with carefully prechosen operators that are spatially related sets of geometrically related elements of the input set. Integral geometry, as suggested by Novikoff (1960), appears to call for operators that are similarly carefully prechosen sets of sets of operators that are then randomly sampled within each set. The methods without

careful prechoice (such as Rosenblatt's) show, in their relatively poor results, the inadequacy of weak sets of operators within so large a space. The methods with prechoice seem too specific to the problem, too unrealistic, and too fragile in the face of variations on the problem. The present method is an attempt to solve the problem described by Selfridge and Neisser (1960, p. 68): "The most important learning process of all is still untouched: No current program can generate test features of its own. The effectiveness of all of them is forever restricted by the ingenuity or arbitrariness of their programmers."

LANGUAGE LEARNING THROUGH RECOGNITION OF REDUNDANT PATTERNS

In a character recognition program, the essential problem is that of grouping together individual examples of patterns into their appropriate pattern class. If the program operates as a continuous learning process, then the procedure is first for the program to try to identify the particular input pattern as an example of some pattern class about which it has already stored information from previous examples, and then for the program to be informed of the actual class to which it belongs, e.g., *A*, *B*, or *C*. This second step provides the feedback information by which the program learns and is able to improve its recognition ability.

The question arises as to whether this grouping of examples into a class and assigning this class a name or code, which is now done by an agent external to the program, could not be replaced by a part of the program itself.

It is certain that a program whose input consisted solely of bit matrices representing the visual images of different characters could never come up with the computer code which represents the name for one of these characters. But would it not be possible for the program to group these characters together into their classes and to reply in its input language as to the class of a pattern by perhaps producing the bit matrix of another, perhaps more typical, member of the same class?

In considering human language this question becomes increasingly important because now the input language of the program, e.g., English, is an ideal output language. Indeed a program which operated in this manner would appear to learn under the same circumstances as a child would. To the young child, human language must appear as an unintelligible sequence of sounds, and to the computer program written language at first would seem to be an unintelligible string of characters. Yet the child is able to reply, at first in syllables or words and later on in phrases and complete sentences. It seems reasonable to ask whether there may not be something in language itself which eventually allows this unintelligible sequence of sounds to become meaningful. The answer may lie in the redundancy of language which, to some extent at least, provides the means for the language's own interpretation.

In written language, patterns exist on many levels, ranging from the simple combinations of letters (TH, TION), to phrases (THE —— ARE ——,), and eventually to the format of whole sentences or paragraphs.

Let us consider, then, a learning program to which a large amount of written language has been given, but which has no explicit feedback from outside to direct its learning process. The program would have a built-in ability to identify patterns as being similar. For example, patterns might be considered which are of a form having alternating fixed character strings and strings of characters which vary from one example of the pattern to another. Examples of patterns would then be considered to belong to the same class if the fixed strings of the examples were identical and the variable strings were of approximately the same character length. As an illustration the two pattern examples: $z(x + y) = z*x + z*y$ and $m(p + o) = m*p + m*o$ would belong to a pattern class:

$$-(- + -) = -*- + -*-.$$

Having identified different patterns, the program might then learn the relationships which hold between different variable strings in a single pattern. In the above example, each variable string is a single character and the relationships which hold are simple. The

fourth variable equals the first variable, the fifth equals the second, the sixth equals the first, and the last variable equals the third one. If now the program encounters: $a(b + c) =$ it can produce the reply: $a*b + a*c$. It has only to recognize the input as the beginning of a known pattern and then simply supply the fixed strings of the pattern along with the appropriate variables according to the relationships it has learned to expect between variables in this pattern.

The program can now reply in the language of the input. The portion of the program which learns the relationships of variables within a pattern depends upon an implicit feedback (telling it what patterns should be similar) provided by the language itself as analyzed with the program's rules for similarity.

Admittedly, the learning of algebraic manipulation shown above is quite simple. But once the language has been broken down by the program into a number of patterns, the relationships which can be determined between variables in a single pattern may be complex, since these relationships may then be defined in terms of the occurrence of the variables in other patterns in a large variety of different ways. Additionally, the program can look for larger patterns defined in terms of a number of smaller patterns, and here too the relationships may be of a complex order.

Having read and analyzed some language, the replies of a learning program capable of this sort of complex association would certainly be unpredictable. Whether they would eventually be as subtle or as intelligent as the replies of a child will probably not be known for many years. While there are still many problems to be solved, it would seem that a fairly simple program might be written which would demonstrate the feasibility of this approach of pattern recognition in language.

GENERAL PRINCIPLES FOR ADAPTIVE COMPUTERS

The specific problems presented to the computer themselves define an input and an output space. A specific problem poses the computer the task of operating upon a specific state (the input state, the hypotheses, the premises, the questions) of this space in order to transform it into a second specific state (the output state, the conclusion, the theorem to be proved, the answer). Both of these states lie within the same space, and the necessary operators can be discovered and defined in terms of, and therefore are composable from information about, states of this space. Operators are simply arbitrary transformation rules about the manner in which states of the space can legitimately change into other states of this space. This parallels, for digital computer programs, possibilities that MacKay (1956) has discussed for adaptive filters.

A computer without any built-in operators can, then, to the extent that it is able to sense (is "aware of") the patterns in the states of the space upon which it must operate, compose its own operators out of these same sensed states. That is, the input sensing organs themselves form the atomic elements for operators, and the states into which sensed inputs (for example, problems, percepts) throw the computer present to the meta-computer that will form the specific computer all the information that it needs (given unlimited time), and as much as it can get, as to how to compose these sensing organ elements into operators. Note that the output, response or answering, elements must also be located within this same space; otherwise, the problem will be unanswerable. And no computer can operate outside of the input space of which it is aware. The process through which input transducing operators are themselves combined into more powerful and efficient operators becomes a process of feedback and reciprocal use of information between the input and subsequent operator formation stages. The operator formation stage is thus a pattern recognition procedure. Its outputs themselves can be either the desired answer or the inputs for a subsequent pattern recognition transformation.

The problem, then, in the designing of such a computer, becomes one of specifying (or, again, treating this as a problem of learning-through-pattern-recognition to map input states appropriately, having the computer learn to specify) methods for analyzing input, operator and output spaces, in order

to generate appropriate operators and to assess whether these operators can be reasonably expected to span the gap between input and output; and, when they cannot, to partition the problem into reasonable parts, to be spanned by a series of sets of operators, whose inputs are the outputs of the immediately preceding set.

At various steps, the computer, when it acts as a meta-computer that assesses the problem space involved, is doing much what Minsky (1959) and especially Newell, Shaw and Simon (1958) have discussed as "heuristic thinking." In particular, the assessment of the imput-operator-output space is similar to the Newell, Shaw and Simon "General Problem Solver's" built-in heuristics of estimating distances and developing abstract schemes for solving problems (1960). But the present conceptualization would embody these procedures quite naturally into the general flow of operation of such a computer, seems more general and complete in its workings, and does not call upon a gigantic specially built-in mechanism of sophisticated programs to effect its purposes. In passing, it seems likely that the collection of things that have been termed "heuristic methods" might best be understood as actions of the meta-computer, which is that part of the computer that (sometimes explicitly, sometimes implicitly) gathers information about some other part, and serves to control this other part as a function of the information gathered.

A SELF-ADAPTIVE NEURAL NET

The last model, one with several immediate neurophysiological interpretations, would seem to suggest a rather "natural" way of both composing suitably constrained and hence sufficiently powerful operators as a function of information learned about the problem space and suitably partitioning the transformation space. This is a layer of parallel input elements that feeds directly back into a second layer of similar elements via "exclusive or" differencing operators (or, functionally equivalent alternatives, Milner's [1958] feedback-transmission net, or a network of excitatory and inhibitor connections, such as those analyzed by Rashevsky [1948] and Nieder [1960]). The "exclusive or" operator is one that fires only when the things it is comparing are not alike with respect to the aspects it senses. This transformation of input into output computes differences, in whatever input set is effectively sensed by the input operators (that is, in the input-as-known), to a precision that is limited by the sensitivity and reliability of the net. There is a great deal of neurophysiological and psychological evidence (see discussions by Horvath, 1960; Uhr, 1959; Miller, 1956) that this precision, in the human computer, is low—on the order of seven discriminations along any single dimension. This sets a limit on the reduction in space size that one layer of these operators can effect, and puts such a net in the position of being able to improve in performance only by means of the addition of new layers of nets. Thus several of these nets of parallel elements layered into series might act as follows: first, define an input space (which is, simply, the space of inputs-as-sensed by the first layer of the net); second, compute differences on this space and construct, by scaling and factor-analytic like methods, the scales and dimensions of an n-dimensional representation of this space; third, assess and then compare local geometric characteristics by computing differences, differences between differences, and higher order differences as necessary, along the dimensions of this space.

For visual pattern recognition, this net would be constructing a space with the dimensions of the visual world—the three spatial dimensions, intensity, and color dimensions. It would look much like the retina: parallel units that feed back in a spatially determined but not completely specified manner, through several series of transformations (synapses), to give, as their output, patterns of numbers (or neural firings) that are functions of the geometric characteristics (such as angles and curves) and intensity gradients of the input, characteristics of the sort Hubel and Wiesel (1959) have discovered are computed in the cat, and Lettvin, Maturana, McCulloch and Pitts (1959) have discovered in the frog. Thus the two or three layers of the retina (and, possibly, additional layers for similar

serial transformations in the cortex) suggest methods for building up transformation operations, by means of a few simple steps in series, that may be sufficiently powerful to map effectively the enormous sets of potential inputs with which the computer must deal. This, then, is a sufficiently shallow depth of computation to be plausible physiologically and to adapt itself and operate under the control of feedback loops. But, in contrast to the single layer of randomly connected neural unit operators (as are, essentially, the building blocks of proposals such as those tested by Rosenblatt [1958] and by Rochester, Holland, Haibt, and Duda [1956]), they would seem to take into account more of the known specificity of connections in the nervous system (at least in the retina) and to give promise of formation of operators with some hope of success.

For temporal or auditory pattern recognition, the net would similarly operate on differences within these modalities as sensed by it. In doing mathematics, this net could learn to compute lengths, slopes, and curvatures (as in fact it does when it operates upon visual patterns) and thus do geometry. It could learn, by differencing among the integers, to order, to count, and to develop concepts of numbered and ordered sets. It is, in general, a net that discovers, compares, and tries to jump across gaps.

DISCUSSION

The suggestion, then, is that the only space within which any computer can operate is the space defined by its own input elements and the inputs presented to them; that is, by the input world as sensed. And only when the desired output also lies within this space is there even a possibility that the computer can succeed at its task of transforming input into output. But information as to input states and output states can be sufficient guide to the pattern-recognition computer in composing, from the elements of the problem space, a set of transformation operators. The chief difficulty lies in the size of the problem space in relation to the amount of information that can be gathered about it in the normal course of operations. The successive layers of parallel units that appear, from anatomical evidence, to be present in the layered neural nets of, for example, the retina and the cortex, suggest the partitioning of this operation procedure into smaller steps, each of which operates upon a space small enough to be manageable. Built-in *im*precision of operations (as seen in the nervous system) itself can be a means whereby the space can be reduced, to the number of just-noticeable-differences (or significant digits) that can be differentiated by the operator. Additional layers in nets of this sort can themselves be differentiated by means of random variations in connectivities within nets, for some of these variations will be looping *within* a layer, which are, in effect, rudiments from which new layers can be evolved by natural selection procedures. Finally, it is suggested that, once such a computer has, through sufficient experience, attained a certain level of ability to dimensionalize and operate within the space of which it is aware, the assignment of names to its operators and especially to its output sets—that is, the teaching of words and, later, mathematical symbols, with which to operate upon and manipulate its "concepts"—would be an extremely plausible means of endowing it with more powerful tools with which it could now stabilize, sharpen, and clarify a conceptualizing system whose adaptive strength was rooted in its operational imprecision.

REFERENCES

Bledsoe, W. W., & Browning, I. Pattern recognition and reading by machine. *Proc. West. Joint. Comp. Conf.*, 1959, 225–232.

Friedberg, R. M. A learning machine: Part I. *IBM J. Res. Develpm.*, 1958, 1, 2–13.

Gelernter, H. Realization of a geometry theorem-proving machine. *Int. Conf. Information Proc.* Paris: UNESCO, 1960. Pp. 273–282.

Horvath, W. J. A neural model for information processing. *MHRI Memorandum*, 1960.

Hubel, D. H., & Wiesel, T. N. Receptive fields of single neurons in the cat's striate cortex. *J. Physiol.*, 1959, 148, 574–591.

Kilburn, T., Grimsdale, R. L., & Sumner, F. H. Experiments in machine learning and thinking. *Int. Conf. Information Proc.* Paris: UNESCO, 1960. Pp. 303–309.

Lettvin, J. Y., Maturana, H. R., McCulloch, W. S., & Pitts, W. H. What the frog's

eye tells the frog's brain. *Proc. IRE*, 1959, 47, 1940–1951.

Mackay, D. M. The epistemological problem for automata. In C. E. Shannon and J. McCarthy (Eds.), *Automata studies*. Princeton: Princeton University Press, 1956. Pp. 235–252.

Miller, G. A. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.*, 1956, 63, 81–96.

Milner, P. M. Sensory transmission mechanisms. *Canad. J. Psychol.*, 1958, 12, 149–158.

Minsky, M. L. Artificial intelligence and heuristic programming. *Mechanisation of thought processes*. London: H. M. S. O., 1959. Pp. 5–27.

Newell, A., Shaw, J. C., & Simon, H. A. The processes of creative thinking. *P-1320*, The RAND Corporation, Santa Monica, California, 1958.

Newell, A., Shaw, J. C., & Simon, H. A. Report on a general problem-solving programme. *Int. Conf. Information Proc.* Paris: UNESCO, 1960. Pp. 256–264.

Nieder, P. Statistical codes for geometric figures. *Science*, 1960, 131, 934–935.

Novikoff, A. Integral geometry, an approach to the problem of abstraction. Paper read at *Bionics Symposium*, Dayton, Ohio, 1960.

Rashevsky, N. *Mathematical biophysics*. (2nd ed.) Chicago: University of Chicago Press, 1948.

Rochester, H., Holland, J. H., Haibt, L. H., & Duda, W. L. Tests on a cell assembly theory of the action of the brain, using a large digital computer. *IRE Trans. Information Theory*, IT-2, 1956.

Rosenblatt, F. *The perceptron. A theory of statistical separability in cognitive systems*. Buffalo: Cornell Aeronautical Laboratory, Inc. Report No. VG-1196-G-1, 1958.

Samuel, A. L. Some studies in machine learning, using the game of checkers. *IBM J. Res. Develpm.*, 1959, 3, 210–229.

Selfridge, O. G., & Neisser, U. Pattern recognition by machines and men. *Sci. Amer.*, 1960, 203, 60–68.

Uhr, L. Machine perception of printed and handwritten forms by means of procedures for assessing and recognizing Gestalts. Paper presented at *ACM Meeting*, Boston, 1959.

(Manuscript received October 25, 1960.)

COMPUTER PROGRAM ABSTRACTS

Multiple Scalogram Analysis for MISTIC, ILLIAC, SILLIAC, and CYCLONE, *James C. Lingoes, The University of Michigan* (**CPA 35**)

This routine resolves a binary measurement matrix into a number of submatrices, each of which fulfills Guttman's criteria for unidimensional scales. This program is eminently suited for the analysis of qualitative data but may also be used for dichotomized quantitative variables. The method (Lingoes, 1960) is a non-parametric data-reduction technique for maximizing inter-item homogeneity based on set-theoretic principles. It differs chiefly from Guttman's scaling technique in its multiple and empirical approach to a set of items, where items are permitted to concatenate as they will without imposing any logical restrictions on the items to be included.

The present program has been written for a 1,024 Williams memory with card input (IBM 528), but can easily be modified for tape input and 4,096 words of core storage. The output is on teletype tape. If k is the number of items; n is the number of persons; and Δ is the number of words required for the responses to each of the items (where $\Delta = n/39$, rounded up to the next integer if fractional), the capacity of the program is given by the following inequality: $k(\Delta + 1) + n \leqq 736$.

A complete description of the method can be obtained from the Computer Laboratory, Michigan State University, East Lansing, Michigan, by requesting Library Program K9M (*Computational Report, I*, 1960).

Reference: Lingoes, J. C. Multiple scalogram analysis. A generalization of Guttman's scale analysis. Unpublished doctoral thesis, Michigan State University, 1960.

An IBM 650 Program for Factor Estimation, *F. J. King, Florida State University.* (**CPA 36**)

Description: Given a rotated factor matrix (F) and a non-symmetric matrix of correlations (R) between the tests in the factor matrix (j) and a set of "new" tests not in the factor matrix (t), estimate the loadings of the "new" tests on the existing factors. The program utilizes the formula $V_{tm} = R_{tj}F_{jm}(F'_{mj}F_{jm})^{-1}$ where V_m is the matrix of the estimated loadings of the new tests on the rotated factors, and m is the number of factors. The method is outlined by Fruchter (*Introduction to factor analysis*, D. Van Nostrand, 1954), having been originated by Dwyer (*Psychometrika*, 1937) and elaborated by Mosier (*Psychometrika*, 1938). **Limitations:** The factor matrix may not exceed 40 variables or 10 factors, the correlation matrix may not

contain more than 10 new tests. **Computer:** IBM 650 with index registers, floating point, and core storage. **Running time:** Approximately 35 minutes for a problem of maximum size. **Comments:** Although the program was written in 650 Fortran, it is a general program within the limitations described above. Intermediate result may be punched if desired. **Availability:** A program write up and a set of program cards can be obtained from the Department of Research and Testing, School of Education, Florida State University, Tallahassee, Florida. In addition to the machine language object program, Fortran statements are available if desired.

An IBM 650 Program for Transformation Analysis, *F. J. King, Florida State University* **(CPA 37)**

Description: Given two orthogonally rotated factor matrices (F_1 and F_2) with common tests, the program computes a correlation matrix (T) between the two sets of factors according to the method of Ahmavaara (*Transformation analysis of Factorial Data*, 1954) and Kaiser (*Psychometrika*, 1960). The formula used is $T_{12} = (F_1'F_1)^{-1} F_1'F_2$. Subsequently, T_{12} is normalized by rows. It is not necessary that the number of factors be the same in the two matrices. **Limitations:** Neither of the factor matrices may exceed 30 tests or 11 factors. **Computer:** IBM 650 with index registers, floating point, and core storage. **Running time:** Approximately 45 minutes for a problem of maximum size. **Comments:** The program was written in 650 Fortran; however, it is a general program within the limitations described above. Intermediate results may be punched if desired. **Availability:** A program write up and a set of program cards, Fortran statements as well as the machine language object program if desired, can be obtained from the Department of Research and Testing, School of Education, Florida State University, Tallahassee, Florida.

A Bendix G-150 Program for Evaluation of the Repertory Test, *Harrie F. Hess, University of Oregon Medical School,* and *Charles R. Baer, University of Colorado* **(CPA 38)**

This program facilitates evaluation of Kelly's Repertory Test, a psychological test used in both clinical and research applications. The program computes "generality" scores for each construct entered in a Repertory Test grid up to 19 × 20 in size. It will also scan the grid by figures, thus computing generality scores for each of the persons whose names are used as "elements" in the grid. The program also computes a total "complexity" score for the entire grid. Data entry can be through keyboard or punched tape. Operating time is about five minutes per grid after data have been entered. Copies of the program and instructions for its use may be obtained from the Director, Numerical Analysis Center, University of Colorado, Boulder, Colorado.

Life does not give itself to one who tries to keep all its advantages at once. I have often thought that morality may perhaps consist solely in the courage of making a choice.

LEON BLUM

ABSTRACTS OF CURRENT LITERATURE WITH INTERDISCIPLINARY IMPLICATIONS

Prepared by Members of the Staff of the Mental Health Research Institute

MARSHALL, N. **ESP and memory: A physical theory.** *Brit. J. Phil. Sci.*, 1960, 10, 265–287.

This article is an attempt to bring ESP (extra sensory perception) in particular, and neural behavior in general, within the scope of physics (i.e., quantum mechanics).

Telepathy is defined as a partial reproduction in one brain of a pattern in another. The author formulates an hypothesis of "eidopoic" influence (meaning: creating a form or pattern) analogous to Newton's law of force. Significant deviations from physics occur, according to the author, only under two conditions: (1) complex structures must be involved; (2) the structures must have "indeterminate" features in their observable behavior. The author assumes, as have previous attempts to join inconsistent findings within a science, that inconsistencies can be explained by presuming that parameters are involved in the new laws whose *values* differ widely from the parameter values involved in previous experimentation. The author defines the concepts of *structure*, *pattern* (the form of a structure at any one time), the scalar of *complexity* of a pattern, the scalar of *difference* between two patterns which makes possible the concept of a "pattern space" (of an indefinitely large number of dimensions), and the concept of *resonances* or attractions between points in this space which, like vectors, have magnitude and direction. Eidopoic Influence is made a function of complexity of patterns and distance between points in the pattern space and is then used to compute a *bias* on the quantum mechanical probability function that structures will change from a certain point to a different point in the pattern space over time. This bias would allow structures in a pattern space to become aggregated into groups or describe orbits. The author describes, and then excuses on the grounds of general scientific tactics, some features of the theory which go against the grain, e.g., the postulation of action at a distance, co-ordinated action at many points in a structure, and the postulation that some of the indeterminancies in quantum physics are in fact determined in a wholistic way. To transfer the discussion from physics and brain physiology to psychology (e.g., telepathy, memory, and perception), the assumption is made that similar psychological states correspond to similar patterns of neural activity. Memory, particularly long term memory, is "explained" in terms of resonance influence (not tied to any particular neurones). (*JWG*)

MACLAY, H. AND OSGOOD, C. E. **Hesitation phenomena in spontaneous English speech.** *Word*, 1959, 15, 19–44.

The study of hesitation pauses and self-corrections offers information on the compositional "encoding" units of the speaker. Here the authors present a distributional analysis of pauses and corrections in the taped record of an academic conference, all in all, some 50,000 words of conversation.

Measures of hesitation were: (a) false starts, (b) repeats, (c) filled pauses ("ummmm", "ahhh", etc.), (d) unfilled pauses (silence). Using the Fries system, each word of text was coded by grammatical form class into (a) lexical words: nouns, verbs, adjectives, adverbs; and (b) function words: articles, possessive pronouns, verb auxiliaries, "not," connectives, and prepositions.

False starts tend to occur with lexical words ($p < .01$) while repeats tend to occur with function words ($p < .001$). Retraced corrections of lexical items almost always include the antecedent function word, but retraced corrections of function words seldom include antecedent words. Unfilled pauses are like repeats, whereas filled pauses tend to occur at grammatical boundaries. If

it is assumed that a speaker will retrace to a boundary of an "encoding unit," the results suggest that one kind of unit contains a function word followed by a lexical word.

Lexical and grammatical determinants operate simultaneously in speech. While it is true that false starts and unfilled pauses occur at points of high statistical uncertainty, correction retraces to the grammatical boundary preceding this lexical choice. Of interest is the fact that subject personal pronouns behave on all measures like function words, suggesting that nouns and pronouns may have important functional differences not suggested by their lexical similarity. (*WPL*)

DORFMAN, ROBERT. **Operations research.** *Amer. Econ. Rev.*, 1960, 50, 575–623.

What is new here is that a professor of economics was subsidized by a major foundation to appraise the field of operations research. He reviews the most common models employed, and subjects them to criticism on the basis of economic theory as well as the practical consequences of advice taken in the form of recommendations of operations researchers.

He concedes that these models, and the methods recommended by the profession for choosing appropriate ends or goals, have made a noteworthy contribution to the principles of optimizing value, and suggests that the minority of economists in the ranks of operations researchers, dominated as it is by graduates from physics and engineering, have made an impression upon the methodology far out of proportion to their numbers. (*RLM*)

BOYARSKY, L. L., SANT 'AMBROGIO, G., AND FRAZIER, D. **Evoked slow and unit responses of the cerebral cortex following repetitive stimulation.** *Amer. J. Physiol.*, 1960, 198, 511–514.

Stimulation of the radial nerve in an intact animal results in two types of activity in the sensorimotor cortex. There is a surface-positive primary potential evoked in the contralateral cortex. This appears to be the postsynaptic response of neurons lying beneath the cortical surface which act as sinks for the positive current flow. It is a slowly varying potential of about 10 milliseconds duration. In addition to the slow potential, single cells in the cortex give rise to characteristic spike discharges. The behavior of both types of cortical activity was studied as the rate of stimulation was increased.

The slow evoked potential on the cortical surface decreases rapidly in amplitude as the rate of stimulation is increased. This decrease is noticeable as the stimulation frequency starts exceeding 1–2 per second and the maximum frequency at which any slow response was detectable was about 60 per second. The usual frequency at which slow waves fail to be detectable is about 30 per second.

Single cells continue to fire at much higher rates of stimulation. Typically, the firing of single units becomes intermittent at stimulation rates of about 40 per second; thereafter it is possible to describe their behavior in terms of the probability of firing which gradually decreases as the stimulation rate is increased. A significant probability exists up to stimulation rates of 60 per second and some units have actually responded to frequencies of 500 and 1000 per second. It is always possible in a given animal to find units which will continue to fire when all slow waves have disappeared from a very localized region. On the average, the cut-off frequency for single units is about twice that for the slow waves.

The difference in behavior of single unit spikes and slow wave activity with increasing frequency of stimulation, taken with other evidence, suggests that the slow responses and unit discharges are independent events. Since the slow waves have been extensively used for mapping afferent inflows to the cortex, this suggests caution in interpreting experiments using these potentials for afferent representation in the cortex. (*WJH*)

SCOTT, P. AND WILLIAMS, K. G. **A note on temporal coding as a mechanism in sensory perception.** *Information and Control*, 1959, 2, 380–385.

This paper suggests that the differential transmission times of nerve impulses from peripheral receptors to the brain is used as a method of information coding by the human

nervous system. The experimental evidence for this comes from tests made on the visual system of human subjects. Pairs of flashes in blue, green, and red light, having a duration of a few microseconds, were presented either simultaneously or with a time interval between them. In order for the subjects to report both flashes of light occurring simultaneously, there had to be a time difference between the presentation of the flash stimuli, the latency difference between red and blue flashes being of the order of 10–15 milliseconds. Since electrophysiological studies of an anesthetised cat have shown that the time of transmission of a light stimulus from the eye to the brain is dependent on the wave length of the stimulus, the authors suggest that a similar explanation holds for the present experimental results.

The suggested mechanism for color perception would then be as follows: Colors are detected by some mechanism within the nervous layers of the retina. The fibers of the optic nerve are connected to the retina in such a way that fibers with a given conduction velocity are attached to the appropriate color detection centers. The effect at the brain of impulsive stimulation of the retina would be a temporal pattern which could be related to the spectral characteristics of the visual stimulus. If such a mechanism forms the basis of normal color perception, further apparatus would be necessary to generate discrete stimulation of the retinal receptor cells, otherwise the temporal coding would be lost in continuous stimulation. It is suggested that the continuous, involuntary tremor of the eyeball, which has been shown to be necessary for visual discrimination, may transform continuous visual inputs into sets of discrete stimuli in the sweep from one receptor to the next.

It is conjectured that this model may be applied to other parts of the nervous system. The peripheral musculature, for example, also has an involuntary tremor with a dominant frequency of about 10 cycles per second, about the same as that found in the eyeball and alpha rhythm of the brain. Temporal coding may then be involved in the transmission of the finer and as yet unexplained details of cutaneous perception from the skin to the brain. (*WJH*)

YNGVE, VICTOR, H. **A model and an hypothesis for language structure.** *Proc. Amer. Phil. Soc.*, 1960, 104, 444–466.

A model of sentence production is set up in detail which yields the following picture of language production. Human speech is produced by a finite-state device and by an essentially left-to right process. There exists a temporary or working memory that can contain about seven items. This is used to make possible a factoring of the state, which results in great economies in the use of the permanent memory. This factoring is represented in the model by an "immediate constituent structure" grammar. The actual process of sentence production corresponds in the model to applying rules by successive expansion from the top down. This process of applying constituent structure rules, if unrestrained, would lead to sentences requiring more than the available temporary memory capacity, since a construction must be developed fully before the system passes on to its successor, and a construction may be developed regressively which thwarts this process. On this account, grammars contain various restrictions and devices like postponement transformations to render them effectively well-behaved. Thus the grammar is brought back within the capabilities of a finite-state device.

The speaker is not ordinarily aware of the number of units he has stored in his temporary memory at any one time; the grammar does this counting for him. Whenever speakers embark on grammatical sentences that exceed the depth limit of the temporary memory, they will become trapped and have to stop and start over. If this is frequent, speakers will try to avoid the constructions that got them into trouble. Many of these constructions will be useful in other sentences so they will not be given up, but in time, alternative structures will develop. At first there is a feeling of stylistic inelegance for deep structures and preference for the others. But gradually the grammar of a language will change in such a way that the stylistically poor constructions become ungrammatical, and the alternatives become the rule. (*WPL*)

ABOUT THE AUTHORS

Arthur W. Burks ("Computation, Behavior, and Structure in Fixed and Growing Automata"), professor of philosophy at The University of Michigan, taught electrical engineering at the University of Pennsylvania during the second world war and was one of the principal designers of the first electronic general-purpose digital computer called the ENIAC. He worked with John von Neumann on the Institute for Advanced Study computer, the JONIAC, and has also consulted with the Burroughs Corporation, and with the Argonne National Laboratory on their computer, the ORACLE (for Oak Ridge National Laboratory). He attended DePauw University, where he received his B.A. in mathematics, and The University of Michigan where he obtained an M.A. and Ph.D. in philosophy. Professor Burks, the author of numerous books, reviews, and major articles, edited the last two volumes of the *Collected Papers of Charles Sanders Peirce*. A member of the Editorial Board of *Philosophy of Science*, his current research interests lie in the philosophy of probability and induction, and application of mathematical logic to the theory of automata and digital computers.

Lee D. Cady, Jr. ("The Factor Structure of Variables Concerned with Coronary Artery Disease") is a specialist in preventive medicine and public health and has a major interest in cardiovascular epidemiology. Assistant professor at the Institute of Physical Medicine and Rehabilitation, New York University, and engineering scientist for the College of Engineering at the same school, he received an M.D. degree from Washington University School of Medicine, and a doctor's degree in public health from Yale. A Fellow of the American Board of Preventive Medicine and Public Health, Dr. Cady is currently devoting full time to the development of modern methods for processing data relating to chronic disease epidemiology. Past experience includes service as an epidemiologist evaluating poliomyelitis immunization effectiveness for the Public Health Service; and service as a health officer in Iran, with the International Cooperation Administration, in charge of programs for communicable disease control, malaria control, and for maternal and child health programs.

Jacob B. Chassan ("Stochastic Models of the Single Case as the Basis of Clinical Research Design"), mathematical statistician, Office of Education in the Department of Health, Education, and Welfare, Washington, D. C., attended City College of New York where he obtained his B.S. degree in mathematics. He also attended the Washington School of Psychiatry and George Washington University, receiving a Ph.D. degree in theoretical statistics from the latter institution. The author of numerous articles in psychiatric, mathematical, and other professional journals, his present research and interests include the development of clinical research programs based on a statistical view of individual psychopathology and interpersonal relations; and research on the mathematical properties of stochastic models directly relevant to the field of clinical psychiatry. Dr. Chassan writes, "One of the great difficulties in clinical research today seems to be the rather limiting ideas that many researchers appear to have concerning approaches which can be considered as statistically or scientifically valid and useful. I am hopeful that my paper will make some contribution toward the broadening of methodological horizons in actual clinical research."

Karl W. Deutsch ("A Note on the Appearance of Wisdom in Large Bureaucratic Organizations"), professor of political science at Yale University, is a member of the editorial board of *Behavioral Science*.

Menard M. Gertler ("The Factor Structure of Variables Concerned with Coronary Artery Disease"), associate professor and director of research at the Institute of Physical Medicine and Rehabilitation, New York University Medical Center, is a specialist in internal medicine with a major interest in the metabolism of heart muscle. A native of Saskatoon, Saskatchewan, and graduate of the University of Saskatchewan with a B.A. degree, he also attended McGill Medical School, where he obtained an M.D.C.M. degree, and an M.Sc. in physiology. Dr. Gertler has written several papers on coronary disease, among them a monograph with co-author Dr. Paul Dudley White, entitled "Coronary Heart Disease in Young Adults." A certified specialist of the American Board of Internal Medicine with a cardiovascular subspecialty, medical director of Sinclair Oil Corporation, and consultant-lecturer at the United States Naval Hospital, St. Albans, New York, he is currently conducting research in the investigation of cardiac enzymology. He writes, "These studies indicate that specific enzyme factors are closely

related to heart disease. The multivariate characterization of the patient with coronary heart disease is making very encouraging progress."

Lida G. Gottsch ("The Factor Structure of Variables Concerned with Coronary Artery Disease") is an instructor at the Institute of Physical Medicine and Rehabilitation, New York University Medical Center. She received special training and experience in the fields of public health, pediatrics, and pathology from the University of Buffalo, where she received an M.D. degree, as an intern at Children's Hospital, San Francisco, as an assistant resident in pathology at University Hospital, New York University, and from The University of Michigan School of Public Health. Dr. Gottsch is currently conducting research in the epidemiology of chronic disease, notably heart disease, and is especially interested in the improvement of medical-statistical communication and workable interrelationships to improve the value of medical data and prevent loss of information. Her most recent publication is a paper written with coauthors Menard M. Gertler, Max A. Woodbury, Paul Dudley White, and Howard A. Rusk entitled "The Candidate for Coronary Heart Disease."

Nathan Kogan ("Aspects of Judgment and Decision Making: Interrelationships and Changes with Age") graduated with an A.B. degree from Lehigh University, obtained his A.M. and Ph.D. degrees from Harvard University, and is currently research associate for the Personality Research Group of the Educational Testing Service in Princeton, New Jersey. Dr. Kogan has written approximately 20 articles published in many psychological journals which are based on his research activities and interests in interpersonal perception, cognitive processes, social attitudes, and the psychological aspects of aging.

William G. Madow ("A Note on the Appearance of Wisdom in Large Bureaucratic Organizations") received his education at Columbia University where he obtained his A.B., M.A., and Ph.D. degrees. He is presently staff scientist at the Stanford Research Institute and professor of statistics at Stanford University. Professor Madow is currently conducting research on sample surveys, statistical and mathematical aspects of behavioral science, and analyses of and experiments on large systems. He is the author of numerous publications which, in recent years, have been largely parts of Stanford Research Institute reports.

John R. Parks ("A Committee Report on Schizophrenic Language"), mathematical analyst in the Applied Mathematics Section of the Research and Engineering Division at Monsanto Chemical Company, attended Case Institute of Technology where he received an M.S. degree in physics. Relative to his professional training, experience, and interests, Mr. Parks writes, "My work involves the application of mathematical and computing techniques to a broad range of business and scientific problems at the Monsanto Chemical Company. My extra-work interest is the application of information theory to problems, including statistics of language." He adds, "My formal training was in mathematical physics. However, my experience has ranged from superintending a forge shop during the war to working at present in operations analysis. The general approach and regimen of my mathematical training has made it possible for me to find interest in a wide variety of subjects: theories of history, economics, the statistics of molecular structure, statistical thermodynamics, computer theory, and now, the statistics of language. In 1952, I was attracted by the beauty and generality of the mathematics of information theory. Perchance, in 1958, I heard some lectures on personality defects given by Dr. A. K. Busch, Clinical Director of St. Louis State Hospital. We discussed the possibilities of information theory in the study of schizophrenia. The possibilities seemed good enough to warrant the formation of a small group of applied mathematicians and psychiatrists to spend a few hours per month studying schizophrenia. This group has called itself 'The Committee on Schizophrenic Language,' since the language area seemed a good place to start." Mr. Parks has published several articles in professional journals.

Philip A. Smith ("Some Problems and Approaches to Automation of Medical Diagnosis"), a previous contributor to *Behavioral Science*, is a psychologist at the System Development Corporation in Santa Monica, California. The author of several articles, among them a recent one on "A factor analytic study of the self concept in psychiatric patients," Dr. Smith has related how his interest in the growing range of digital computer applications was stimulated while he was working as a clinical research psychologist at the Veterans Administration Hospital in Ann Arbor, Michigan. His initial work with a computer at Michigan to process psychiatric research data led to larger research efforts, and to awareness of computer developments in other areas. He adds "At the System Development Corporation, my interest soon broadened to very large, real-time data processing and control systems." Dr. Smith received his B.A. degree from Miami University and his M.S. and Ph.D. degrees from the University of Kentucky.

Leonard Uhr ("Suggestions for Self-Adapting Computer Models of Brain Functions"), research psychologist at the Mental Health Research Institute of The University of Michigan, became interested in the potentialities of computers as a result of previous psychological interests in perception and concept formation. This article is an outgrowth of work which Dr. Uhr has carried out during the past several years and continued at the System Development Corporation (where he met his coauthor, Charles Vossler) in the summer of 1960. In addition to his work with computers, he has been interested in research on the behavioral effects of psychoactive drugs and, with Dr. James G. Miller, is coeditor of the recently published book *Drugs and Behavior*. Dr. Uhr received his undergraduate education at Princeton University, and did graduate work at Johns Hopkins, where he received an M.A. in writing, and The University of Michigan, where he received a Ph.D. in psychology.

Charles M. Vossler ("Suggestions for Self-Adapting Computer Models of Brain Functions"), programming analyst at the System Development Corporation, was graduated with a B. A. degree in mathematics from Cornell University. Relating how he became interested in machine intelligence, Mr. Vossler writes, "I first became acquainted with programming in 1957 when the Air Preheater Corp. of Wellsville, N. Y., acquired a Burrough's E101 computer to aid in the solution of heat transfer and mechanical engineering problems. This led to an interest in larger computers and in 1958 to programming for the IBM 709 at the Douglas Aircraft Co. in Santa Monica, Calif. Here I had an opportunity to participate in the writing of a 709 program to simulate the operation of the E101. This work on the E101 Simulator greatly stimulated my interest in machine intelligence, because here the fixed code of the simulator could operate a variety of secondary codes; in fact, any problem which could be expressed in the language of the E101. This led quite naturally to the question of whether an appropriate program could create this flexible secondary language by itself, and in this way learn as the result of experience. In 1959, I began work at the System Development Corporation in Santa Monica and in the last few months have been working with Dr. Uhr on his pattern recognition project."

Michael A. Wallach ("Aspects of Judgment and Decision Making: Interrelationships and Changes with Age"), assistant professor of psychology at Massachusetts Institute of Technology, is the author of approximately 15 articles which have appeared in major psychological journals. These articles are concerned with his research interests in personality and motivation in relation to cognitive processes, cognitive development and change with age, and assessment of personality and thinking through media of aesthetic expression. Additional papers in the same areas of research are currently in press. Professor Wallach obtained an A.B. degree from Swarthmore College and a Ph.D. degree from Harvard University.

Max A. Woodbury ("The Factor Structure of Variables Concerned with Coronary Artery Disease"), described as having "a keen aptitude and training for the mathematical analysis of biologic and economic problems," is research professor of mathematics and research co-ordinator with the Research Division, College of Engineering, at New York University. He attended the University of Utah and continued his studies at the University of Michigan where he obtained his M.S. and Ph.D. degrees and was appointed instructor in mathematics and project director at the Engineering Research Institute. He has also held positions as research associate at Princeton University, associate professor and research investigator for "Project Big Ben" at the University of Pennsylvania, and was principal investigator for the Logistics Research Project at George Washington University. Professor Woodbury, whose particular interests are in the development and application of mathematical and computing techniques to problems of time series, pattern estimation, and complex causation, is the author of several articles in these interest areas.

'Logic! Good Gracious! What rubbish!' exclaimed E. M. Foster's Old Lady. 'How can I tell what I think till I see what I say?'

JOHN JAY CHAPMAN, *Emerson Sixty Years Later*

serving *a critical function* in psychology . . .

THE PSYCHOLOGICAL RECORD

a quarterly journal of

theoretical and experimental articles reflecting
a wide range of reader interest, new approaches,
critiques and commentary on current developments.
Regular subscription, $4.00 per year.
American Psychological Association member, $3.00.
Student rate, $1.50.

Mail subscriptions to:

THE
PSYCHOLOGICAL
RECORD

Denison University, Granville, Ohio